THE DEVELOPMENT OF ALLOCATION METHODOLOGY FOR SYSTEM RELIABILITY REQUIREMENTS

A THESIS

Presented to

The Faculty of the Graduate Division

by

Nam Kee Lee

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy in the

School of Industrial Engineering

Georgia Institute of Technology

August, 1966

THE DEVELOPMENT OF ALLOCATION METHODOLOGY

FOR SYSTEM RELIABILITY REQUIREMENTS

Approved:

Chairman

Date approved by Chairman: September 3, 1966

ACKNOWLEDGMENTS

I would like to thank those who helped me in the preparation of this thesis.

In Particular, I am grateful to Dr. William W. Hines for inspiring my interest and serving as the chairman of the committee. The work from which this thesis evolved could not have been done without his constant advice and encouragement.

Many thanks are due to Dr. David E. Fyffe for his guidance and helpful suggestions during the preparation of the thesis. I am also thankful to Dr. Joseph L. Hammond, Jr. for his careful review of the final manuscript.

My sincere appreciation goes to Dr. Robert N. Lehrer, Director of the School of Industrial Engineering, for his understanding and assistance during my study at the Georgia Institute of Technology.

Last but not the least, I am indebted to my wife, Youn-Joo, for her patience and sacrifice throughout the difficult parts of my graduate study. To her I express my special thanks.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

SUMMARY

In any system design, reliability requirements cannot always be met simply by designing a system that will perform the intended system functions. The system designer responsible for designing a reliable system must therefore translate the overall system reliability requirement into numerical requirements at various subsystems and units which comprise a given system. The purpose of this research is to develop comprehensive methods for allocating a system reliability requirement to the various subsystems and lower levels of a system when it is desired to achieve the specified reliability goal consistent with a set of system constraints. Specific areas of investigation included the following: (1) the development of methods for determining the reliability requirements at various subsystem levels within a system which will minimize costs of achieving the specified level of system reliability consistent with a given set of system constraints, (2) the development of a method for allocating a system reliability requirement into subsystem levels based on the concept of total system costs, (3) the development of appropriate computational procedures for the reliability allocation and optimal redundancy problems, and (4) finally the evaluation of reliability improvement obtainable through repairable redundancy at the weak system elements.

Because of the complexity in relating many interacting factors involved in the reliability allocation process, the system reliability allocation is carried out in this research through two distinct methods--

the functional allocation method, and the detailed allocation method. With the application of the functional allocation method, the system designer can assign a first set of numerical requirements for each of the major functional equipment groups in a given system. The supplementary allocations among various subsystems and lower levels of a system are then accomplished by the use of the detailed allocation method.

The development of the functional allocation method was concerned with two fundamental aspects of the reliability allocation process--a mathematical model for the allocation, and the standard input data to the model. In the first area, the objective was to develop a realistic and practical allocation model which is generally applicable to various types of systems. In order to achieve this objective, a mathematical model was developed in which the system reliability requirement is apportioned among the several functional equipment groups in consideration of such factors as unit state-of-the-art, unit complexity, unit duty cycle, unit essentiality to the system success, unit applicational stresses, use environmental factors as well as the overall feasibility of the specified system requirement. The standard input data required for implementing the allocation method as well as the step-by-step procedure for determining them were described in detail. In the functional allocation, the systems which include redundancy were not treated because of the consideration that a preliminary system design in the early phases of a development program primarily determines the type and minimum set of functional equipment required to perform the intended system functions, and therefore the use of redundancy is not a standard design practice at this stage of the system design.

Once the reliability requirements at the functional equipment level have been established, it is the time that the system designer carefully reviews and modifies, if necessary, the allocated requirements as soon as the detailed feasibility study discloses the discrepancies between the allocated requirements and improvement feasibility. The most interesting case arises when there are several design approaches available for achieving the allocated requirements including such methods as (1) redundancy at the weak units versus design simplification or further derating of the system elements, or (2) redesigning weak units versus a search for high reliability components, or (3) combination of these approaches. Any choice of available design approaches to achieve the desired level of reliability must then be based on the sound trade-off analysis which serves to select one optimum design approach that can be achieved for a given expenditure of efforts, usually within the limitations of available funds, system weight, time or manpower.

The detailed allocatin method developed for accomplishing these supplementary allocations takes explicitly account of this existence of several design approaches at each functional equipment group. The allocation problem formulated is then that of selecting the design alternative and the level of redundancy at each subsystem in such a way as to maximize the overall system reliability when (1) there are several design approaches available, and (2) there also exist the constraints on system costs and weight for a given system. An application of dynamic programming techniques to this allocation problem resulted in a modified formulation of the problem in which the constraint on weight is eliminated by the use of the Lagrange multiplier method. It was found that the reduction in dimensionality of the problem through the Lagrange Multiplier method results in a considerable savings in the computational efforts. However, it was also found that the presence of two or more system constraints does not change the basic nature of the reliability allocation problem considered herein.

The step-by-step computational procedure for the allocation problem was described with two hypothetical problems. A computational flow chart that may be used for the computer program was also presented. The nature of the computational procedures developed is such that whenever the modified allocation problem is solved, a by-product of the calculation of the optimal policy is a set of state values of the eliminated constraint function that will permit the sensitivity analysis in the given circumstances. In most cases of the system reliability allocations, this sensitivity analysis is more interesting and useful than solving the problem for a particular value of the constraint. Even in the complex systems containing many system elements, no serious computational difficulties arise since the computational scheme developed is rather simple, practical, and yet yields an exact solution to the problem.

One important concept which is not incorporated in the development of the functional allocation method is that of total system costs which is of prime importance in any development program. A reliability allocation method based on the concept of total system costs was developed in which the allowable system failure rate is apportioned among subsystem levels so as to minimize the sum of the amortized initial investment costs, annual operation and maintenance costs, and the costs associated with the failure of the system. This allocation method has several characteristics. It takes into account the prime cost elements that are relevant in the reliability allocation; and it makes distinction between the one-time initial investments and the periodic recurring costs which will not only permit an easy measurement of relevant cost elements but also actually minimize the possible errors in the estimation of the cost elements. The computational method presented is rather simple so that the solutions are easy to obtain in a straightforward manner. This method was applied to a hypothetical problem in order to illustrate the method of solutions. This allocation method also provides the system designer with a basis for evaluating alternative system designs in terms of both the cost of a given level of system reliability and the system reliability achievable for a fixed amount of funds. The method is thus "mixed" in a sense that it involves allocations of system reliability requirements as well as planning decisions.

The final part of this research was concerned with the evaluation of relative improvement in reliability of the redundant system where some system elements are made repairable while in operation by an attending operator. In certain situations, the system designer can increase reliability of his system by making the weak system elements repairable rather than the straight duplication of them. It is then desirable to find the relative improvement in reliability or in its mean time to failure of the repairable system over its non-repairable counterpart. The reliability functions and the formulas for the mean time to failure were derived for both the repairable and non-repairable systems which will in turn permit to estimate this relative improvement in a given situation. This measure of reliability improvement will lead to optimum methods of choosing on proper combinations between the straight redundancy and the repair policy based on such measures of effectiveness as increased system weight and system costs in any design situation where the allocated requirements exceed the current state of the art capability for a given system.

CHAPTER I

INTRODUCTION

General Concepts

One of the most important and challenging engineering problems in this decade is the design and development of highly reliable systems for both military and commercial uses. Only a few years ago reliability was a qualitative and rather abstract, equivocal term even for those who were engaged in the engineering professions. The lack of understanding of concepts of reliability and the unavailability of much needed techniques of reliability have not only cost billions of dollars but also slowed technological progress in many fields.

Although reliability has long been considered as one of the desired qualities of a product, in the past no adequate techniques existed to achieve this desired quality. However, in the last few years, a large body of methods of reliability analysis has been developed in response to modern needs, making possible sound quantitative analysis of reliability for large complex systems. Reliability is now recognized by industry as one of the essential engineering technologies.

With the rapid technological innovations in all engineering fields, functional roles and mission requirements for a modern system become increasingly sophisticated and complicated. Consequently, engineering systems inevitably become more complex in the functional configurations necessary to satisfy these increased performance requirements. Reliability has entered a new era with the advent of the electronic age in which the equipment analyzes a bulk of business data, makes routine decisions based on the input data, controls automated chemical processes, and provides a communication network or early warning of an enemy's surprise attack. It is not surprising that the number of electronic components used in the modern weapons systems, for instance, has grown from ten to hundreds and even to thousands. Similar trends can also be observed in the design of modern computer systems and control networks used in automated systems.

As system complexity increases rapidly, the concept of system reliability invariably becomes a critical problem. Since the complexity of the system is one of the specific causes of its unreliability, industry has become conscious of the importance of reliability and has initiated the reliability-oriented planning, design, and testing programs in the development of a new system. Reliability has become a primary concern, especially in the design, procurement, production, operation, and maintenance phases of the system development for commercial, military, and space exploration uses.

The concept of reliability is extremely important in the design of systems for military uses, and will be critical in the manned space flight and exploration where failure-free operation is vital to the preservation of human life and highly expensive equipment due to the fact that the system failure usually allows no recovery of the equipment and human life involved. Reliability is of utmost importance to the system designer, especially when failure of the equipment he designs may cause loss of the entire system in which it is used as, for instance, in the manned aircraft systems where loss of life is involved or when it is part of the weapons systems itself which may be a vital link in the security and defense of the nation.

Much of the early reliability work was done by the Armed Services for their weapons systems, and military authorities have since been active in the development of modern reliability techniques. The concept of reliability is of interest to the military establishments since any inadvertent system failure may cause the loss of a military mission and lead to disastrous consequences.

Another reason for the military emphasis on the reliability of the weapons systems is because of high system maintenance costs and the shortage of skilled maintenance technicians. According to one source,[1] maintenance expenditures of the Department of Defense have been estimated at more than $25 million per day, including the employment of about 800,000 military and civilian technicians who are directly engaged in the maintenance of the military weapons systems--a total of about $8 billion per year or more than 25 per cent of the total defense budget. It is also reported that for a sample of four pieces of equipment in each of three classes--radar, communication, and navigation--the annual support cost is 0.6, 12, and 6 times, respectively, the cost

---

[1]Col. V. J. Bracha, *Analysis of Reliability Management in Defense Industries*, USAF Ballistic System Division (BSD-62-48).

of the original equipment.[2]

Military reliability considerations became increasingly important during the post-1950 period. During the Korean War the rapidly expanding use of electronic weapons systems coupled with their increasing complexity and inability to withstand severe military environments under which weapon systems are expected to operate satisfactorily further culminated in the emphasis on reliability.

In 1952 the Department of Defense established the Advisory Group on Reliability in Electronic Equipment (AGREE), which consists of nine special task groups. The primary purpose of the AGREE group was to monitor and stimulate interest in reliability matters and recommend measures which would result in more reliable electronic equipment. The task groups composed of industrial and governmental representatives, published a report in June, 1957, with many recommendations.[3] Their important work has stimulated the growth of modern reliability activities, and many of their recommendations presently appear in the military specifications for electronic equipment and systems. The primary recommendation by the AGREE group was to establish numerical reliability requirements and the necessary program control techniques to properly execute these reliability requirements for all military weapons systems.

---

[2]J. A. Cafaro, and H. D. Voegtlen, "The Measurement and Specification of Product Abilities," *Industrial Quality Control*, Vol. XVIII, No. 9 (March, 1962), pp. 20-26.

[3]Advisory Group on Reliability in Electronic Equipment, *Reliability of Military Electronic Equipment* (Washington, D.C.: U. S. Government Printing Office, June, 1957).

During the past decade, the manufacturer engaged in the design and development of new systems was faced with the necessity of meeting certain performance requirements. Theoretically, he was considered equally responsible for assuring a high level of reliability in the new system. Since the performance requirements have always been spelled out in the contractual specifications in numerical terms, the manufacturer usually places more emphasis upon these requirements in order to fulfill his legal obligation. However, as far as the manufacturer's responsibility with regard to the concept of reliability is concerned, it is generally treated as a secondary requirement in contrast with primary performance requirements. After a design has attained other performance requirements which are rigorously specified, the concept of reliability is then considered as an afterthought, based on the classical engineering statement "we will do the best we can in our product as far as the reliability is concerned."

Engineering experience has shown that this "afterthought" approach toward reliability requirements is not sufficient and bears the costly consequences. The resulting design of a system creates an irreducible failure rate which cannot be cured or eliminated from the finished equipment. When the equipment has been finished, the scheduled development time for the system has already been exhausted and the system is now needed for tactical use. Should the system fail when needed, the money already spent is too great an investment to be written off because of poor reliability.

Very little can be done about unreliability once a piece of equipment has been designed and is in production. Sometimes, components of higher capacity rating can be substituted in an effort to reduce the failure rate of the equipment. However, if a piece of equipment has turned out to be unreliable because sufficient considerations were not given to reliability during the initial design stages, only extensive redesign or complete abandonment of the development program can help to solve the problem, and thus money and time are lost.

In order to avoid undesirable consequences due to unreliability of the system, the system manufacturer must have a sound appreciation for and a comprehensive knowledge of the concept of reliability. Frequent failures in the commercial products of a company due to unreliability will soon cause customer dissatisfaction, lead to loss of prestige, and cause eventual loss of business in the competitive markets.

As science and engineering technologies advance to such an extent that it will be nearly possible to create engineering systems and equipment which once existed only within the realms of scientific imagination, the existing knowledge of reliability engineering must also be extended into new horizons in order to cope with many pressing problems. This is simply because of the fact that what is feasible and possible in the engineering system design does not necessarily lead to a reliable system in most cases.

The major causes of unreliability in every system can be traced to the dynamic complexity of system development program and the lack of proper control techniques for preventing unreliability in the initial phases of a system development. The engineering measures available to achieve very high reliability may be in contradiction to other desirable system characteristics such as reduced size, high performance, urgent delivery, and low cost. From the viewpoint of the concept of system effectiveness, the problem of reliability allocation will necessarily result in the trade-off analysis among the important system parameters.

A typical evolutionary process in the development of a new system would reveal a number of interesting steps from the inception of original idea to the acceptable prototype model, and finally to lot-size production. Figure 1 describes a typical system engineering process as viewed from the concept of reliability. The system design begins with an original idea, well-defined operational need, and intended system functions. The weapons system, for instance, would define the enemy's threat to be encountered, the size of attack to be handled, and expected frequency of enemy's surprise attacks. The tactical need for reliability must first be anticipated, and specific reliability plans must then be laid to fulfill the reliability needs. Reliability requirement is then defined and specified formally in the system design specifications which may include availability[4] and maintainability.[5]

With a set of prepared system and equipment specifications, preliminary system design is to be explored with detailed syntheses of system configurations indicating the kinds, quantities, and characteristics of the equipment that will be required to perform intended system functions. This preliminary design is reviewed and revised to

---

[4]The term "availability" is defined to be the probability that at any point in time the system is either operating satisfactorily or ready to be placed in operation on demand when used under stated conditions.

[5]The term "maintainability" is defined as the probability that a failed system is restored to operable condition in a specified down time.

correct deficiencies, if any. For a system to be developed at the minimum possible cost and within a fixed development time, all alternative designs must be considered and evaluated from a system's approach.

During the prototype development phase, a pilot model of the system is built according to the preliminary design. Reliability of the prototype model is evaluated by life testing, and reliability requirement is thus "proved in" by actual demonstration. The design is refined to remedy deficiencies present in the preliminary design stage, such as crude appearance, excessive weight, bulkiness, or unsuitability for mass production.

In lot-size production phase, the system design is finally materialized in a form of mass production. Parts, materials, and production processes are controlled on routine basis by statistical quality control and inspection programs. Reliability is now built into the system, culminating in a successful operational acceptance test by the prospective users.

During the evaluation phase, the completed system is evaluated from the actual usage experiences at the field to determine that the original requirements have been met, and the resulting failure information is fed back into the system on a timely basis. This information feedback process is shown by arrows in Figure 1. This feed-back of usage information is essential for further improvement of the system, and provides a sound guideline in the future development planning of a similar system.

The development of a reliable system involves complicated problems which are usually interrelated. The complexity of the problem

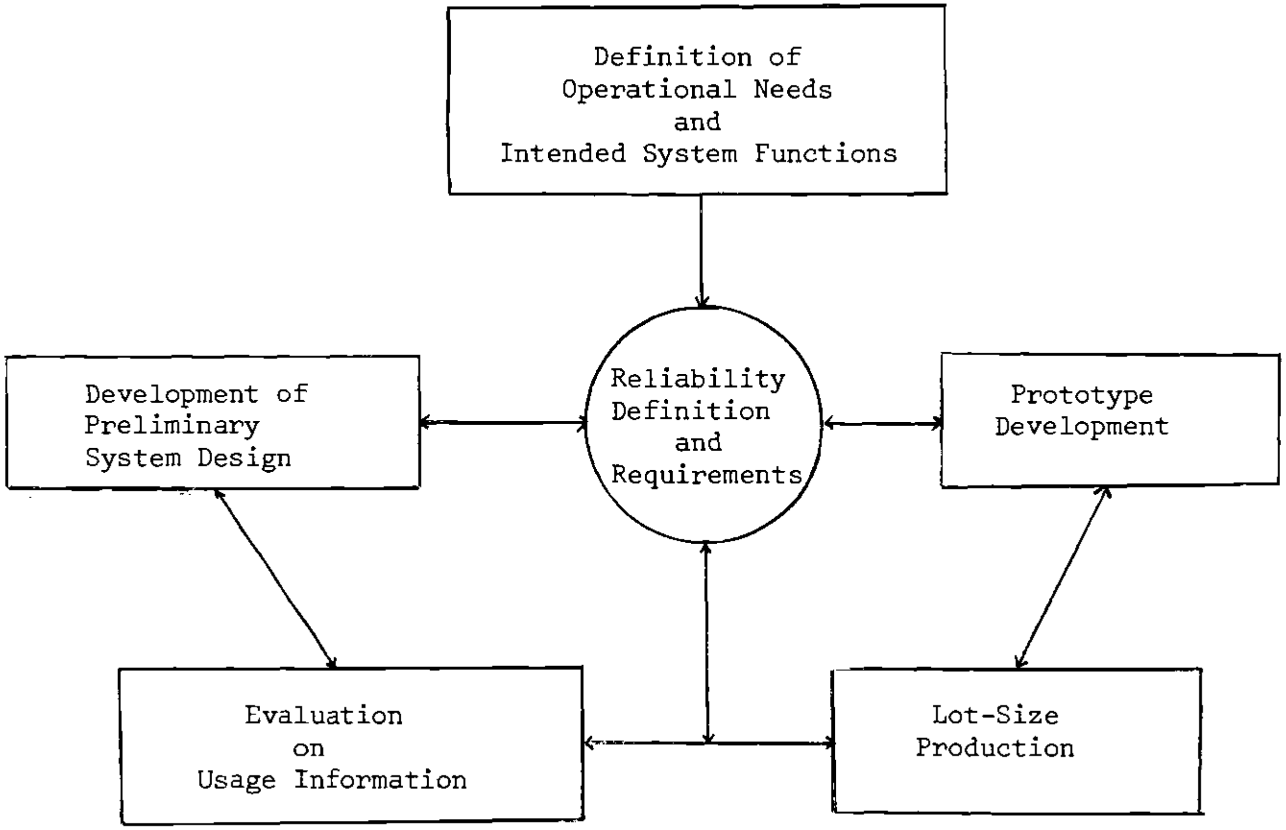


Figure 1. Typical System Engineering Process as Viewed
by the Concept of Reliability

varies with the specified mission of the system, expected adverse consequences of system failure, available knowledge of its feasibility and inherent limitations, and the reliability knowledge gained from similar systems previously developed. In any case, the most important task for the system designer responsible for developing a reliable system is to translate first the overall system reliability requirement into specific numerical reliability requirements at various subsystem levels. Realistic and consistent reliability allocation among the many subsystem levels is an initial means and a basic rule for achieving the desired reliability goal in any system development program.

Nature of the Reliability Allocation Problem

One of the most important problems that system designers as well as project managers face in the early stages of a system design is how to establish quantitative reliability requirements for each of the subsystems which are compatible with their current state of development, expected improvement feasibility, and amount of money budgeted for their development program such that the required system reliability goal is achieved. System reliability requirement is usually established during the initial phases of a system development program by either the end user or proposed by the system contractor on the basis of operational requirements and feasibility.

System reliability requirements may be stated in any appropriate measure such as system mean life, failure rate or probability of successful system operation over a fixed time period. The only basic constraint in the apportionment of a given system reliability requirement is that allocated unit reliability requirements, when recombined, must satisfy the system reliability goal.

In contrast with this single, basic constraint in the allocation, the allocation process involves more than solving a simple mathematical equation with such a single constraint. It generally involves rather complicated and interdependent relationships between unit and system reliability since the units in the system interact with each other so that individual unit reliability varies with the type of function to be performed, complexity of the unit within a system, and operational environments under which the unit is expected to operate.

In developing a realistic reliability allocation model, all factors which influence to some extent unit reliability within an integrated system must be considered if the most economical and yet feasible requirements are to be obtained. The following factors are generally considered important in the allocation of system reliability requirement:

1. *Unit State-of-the-Art:* Current state-of-art capability; the designer's reliability prediction of future improvement feasibility supported by the reliable data sources; results of all life tests conducted to date.

2. *Performance Requirements and Failure Characteristics:* Criteria of success and failure, including degraded performance, tolerance and wear, and parameter shifts; failure modes; underlying failure density function.

3. *Environmental Factors:* Operational stresses under which the unit will be subjected in the use configurations; use environmental

conditions, including storage and transport period.

4. *System Characteristics:* Unit functional importance; unit complexity; adverse consequences of unit failure; system configurations; unit duty cycle; maintenance and checking policy.

5. *Design Alternatives for Reliability Improvement:* Unit derating techniques; redundancy; design simplification; concentrated basic research on the target unit.

6. *Physical Constraints:* Weight, space and power limitations; configuration and packaging requirements.

7. *System Costs:* Cost of achieving the reliability goal; associated system costs including operation and maintenance expenditures; expected costs of system failure.

8. *Evaluation of Trade-Offs between Performance, Maintainability, Weight, Space, Power, Time, and Costs.*

Ideally the most logical approach to allocating system reliability requirement to subsystem levels is to base the allocation on all the factors which will influence unit reliability. However, any mathematical allocation model that includes all of the above factors in a rigorous manner would be extremely difficult to formulate, and virtually impossible to solve in practice. Furthermore it is equally difficult to express each of the above factors in numerical or mathematical terms.

On the other hand, the allocation model which is simplified to the extent that solutions obtained from the application of the model are practically unrealistic and unattainable is, of course, an undesirable approach. Therefore, these factors must be reduced to those elements absolutely essential to avoid excessive model sophistication, yet yielding the realistic solutions.

Although system designers generally have recognized the various factors which will affect unit reliability, they have had no adequate quantitative method of relating the various factors. The problem of realistic allocation is made more complex by the fact that detailed information on many of these factors is not readily available early in the system design stages. Consequently, reliability requirements for the components, units, and subsystems were often determined solely on the basis of past failure data of comparable units which were arbitrarily adjusted to satisfy the basic constraints in the allocation process.

The system designers and project managers responsible for preparing specifications have also had little guidance in establishing realistic reliability allocations by a practical, non-subjective method which takes account of those essential factors affecting the system reliability and produces realistic allocations to be used in the contractual documents. Although several allocation methods are presently used in industry, no single technique has yet gained universal acceptance.

The Advantages of Reliability Allocation

In any system development program, reliability is regarded as a desired characteristic that can be either created or destroyed. The creation of a reliable system can only be achieved through sound planning, design, testing, production, and proper feed-back of usage information according to a set of well-managed reliability assurance programs. The degradation of reliability in an otherwise reliable system comes from neglect or ignorance of such an effective reliability-oriented program in any development project.

Above all, reliability built into a newly-developed system must and can begin only with effective and realistic allocations of system reliability requirement among various constituent units. Reliability allocation during the initial phases of a system design is then the most essential and valuable tool by which the prime objective of building a reliable system can be accomplished.

Some of the important advantages to be obtained through establishing the reliability allocation program may be summarized as follows:

1. The inclusion of numerical reliability requirements in the contractual specifications will ensure that reliability is treated as a design parameter of equal importance with other system parameters such as performance, weight, and cost.

2. Reliability allocation will provide a basis for initial allocation of requirements and supporting research and development effort among subsystems, and will determine the extent of special research and development support required in specific areas.

3. Reliability allocation will assist in accurate prediction of development costs and time required for a given system development program.

4. Reliability allocation imposes upon the system contractor a legal obligation of meeting the specified requirement, and thus force the contractor to be on the alert, throughout the entire development program, to all discrepancies in reliability which may require

contractor's decisions. This will in turn lead to improved system design, procurement, manufacturing, quality control, and testing programs. This would not only ensure a reliable system but, in the long run, would contribute to the development of a scientific approach toward designing for reliability.

5. Reliability allocation in the early design stages provides a standard for reliability and value assurance activities. It also provides a basis upon which demonstration and acceptance tests can be prepared and costed.

6. Reliability allocation focuses attention on the relationships between component, equipment, and system reliability, providing a good basis for the advisability of reliability and maintainability feasibility study prior to the award of a design and development contract.

7. Reliability allocation will establish a starting point on the probable trade-offs among weight, space, power, other performance requirements, and system development cost that will be required to achieve the allocated reliability requirements.

8. Reliability allocation will serve to integrate reliability requirement into production process, and production and quality control specifications.

9. Reliability allocation at primary subsystem levels will serve to incorporate the reliability requirements in subcontractor and vendor procurement documents.

10. Reliability allocation arrived at by the sound allocation procedures would be more realistic, consistent and economical than

subjective and "rule of thumb" methods. Above all, the cost of implementing such an allocation procedure is negligible as compared with the expected savings of time and money expended in meeting specified reliability goals, after a failure has been encountered while the equipment is in use in the field, a failure which results in speedy initiation of a crash reliability program.

In short, the policy of "invest a bit more and save a lot later" is considered to be the primary objective of any reliability allocation program.

Previous Work on Reliability Allocation

It is now generally recognized that any system development program is not complete by itself if it does not quantitatively specify the system reliability requirements during the early development phases of the program. It is also a generally accepted practice that this reliability requirement be stipulated in the contractual documents along with other important design parameters.

Among others, the military authoriteis were and have been the pioneer in initiating reliability activities. Their studies and reports have spurred the growth of various reliability programs in industry, including design, development, procurement, production, operation, and maintenance.

The reliability allocation methods currently in use in industry or those available in publications vary from fairly simple to rather complicated ones. A number of military specifications, standards, and handbooks are now available which either state or imply the reliability requirements as an important system parameter. The following military specifications and documents generally specify the reliability requirements which have received general acceptance and widespread industry applications:

MIL-STD-441 : Reliability of Military Electronic Equipment (June 20, 1958).

MIL-STD-756 : Reliability of Weapons Systems, Procedures for Prediction and Reporting Prediction (October 3, 1961).

MIL-R-19610 : Reliability of Production Electronic Equipment, General Specification of (September 15, 1956).

MIL-R-22256 : Reliability Requirements for Design of Electronic Equipment or Systems (November 20, 1959).

MIL-R-22732 : Reliability Requirements for Shipboard and Ground Electronic Equipment (March 10, 1961).

MIL-R-22973 : Reliability Index Determination for Avionic Equipment.

MIL-R-23094 : Reliability Assurance for Production Acceptance of Avionic Equipment.

MIL-R-25717A: Reliability Assurance Program for Electronic Equipment (March 9, 1959).

MIL-R-26474 : Reliability Requirements for Production Ground Electronic Equipment (June 10, 1959).

MIL-R-26484 : Reliability Requirements for Development of Electronic Subsystems or Equipment (June 2, 1958).

MIL-R-26667A: Reliability and Longevity Requirement, Electronic Equipment, General Specification for (June 2, 1959).

MIL-R-26674 : Reliability Requirements for Weapons Systems, General Specification (June 18, 1959).

MIL-R-27070 : Reliability Requirements for Development of Ground Electronic Equipment (March 25, 1960).

MIL-R-27542 : Reliability Program Requirements for Aerospace Systems, Subsystems, and Equipment (June 28, 1961).

MIL-HDBK-217: Reliability Stress and Failure Rate Data for Electronic Equipment (August 8, 1962).

NPC-250-1 : (National Aeronautics and Space Administration) Reliability Program Provisions for Space Systems Contractors (May 13, 1963).

Among the specifications cited above, only one specification, MIL-R-26474, specifies for reliability requirements at the subsystem levels, and outlines procedures to determine equipment mean time between failure requirements. The procedure is based on a part class count of the equipment under consideration. Tubes, motors and relays, semiconductors, and other electro-mechanical parts constitute the four part classes considered in the method. This approach is not regarded as an allocation procedure since no considerations were given to the functional requirements of the system.

MIL-STD-441 describes many important factors to be considered in the development, planning, design, and construction of prototype models of new equipment. The standard also provides a simple allocation formula based on the historical failure rates of the constituents of the system in serial system configurations.

MIL-STD-756 provides the procedures for predicting the quantitative reliability of aircraft, missiles, satellites, and other electronic equipment, and their constituent components throughout the entire development phases, in order to detect inherent design weakness and to establish a basis for allocation of reliability requirements to the various components of the system.

MIL-R-22256 outlines procedures which will ensure high inherent reliability before release to production, and prescribes the detailed requirements for the feasibility study, planning of the proposed design, and reliability assessment of the proposed design. MIL-HDBK-217 provides the detailed procedures and failure rate data for the prediction of equipment reliability based on the stress analysis of the parts and components used in the design of the specified system.

MIL-R-23094, MIL-26674, MIL-R-27073, and MIL-R-27542 all specify the requirements that the overall reliability goal should be allocated into the constituent elements of the system with due considerations given to those factors such as complexity, consequences of unit failure, required time of operation, and environmental conditions. However, the specifications do not expressly detail how the allocation is to be performed.

Electronic Industries Association,[6] and Frederic[7] have described the simple allocation methods in which system reliability requirement is allocated into the constituent units in proportion to the ratio of the historical failure rate of the unit to sum of the failure rates of all units within a given system. The procedures are based on unit failure rates which are assumed to be constant. The other necessary assumptions are that unit failures are independent of the other units, and that failure of any unit will result in system failure; i.e., the system is composed of units in series.

The reliability allocation method for the electronic equipment described in the Advisory Group on Reliability of Electronic Equipment (AGREE) report[8] is somewhat more sophisticated in a sense that unit complexity rather than unit failure rate is used as the basis of the allocation, and at the same time unit importance factor which measures the probability of the system failure if the specific unit fails is incorporated in the allocation model. In the AGREE allocation model, the complexity factor of an equipment is defined in terms of modules where a module is an electrone tube, a transistor, or a magnetic amplifier, and its associated circuitry.

---

[6]Electronic Industries Association, "Determination of Permissible Component Part Failure Rates," *Electronic Application Review*, Vol. 4, No. 1 (September, 1956), pp. 10-12.

[7]H. E. Frederic, "A Reliability Allocation Technique," *Proceedings of the Fourth National Symposium on Reliability and Quality Control*, (January, 1958), pp. 314-317.

[8]AGREE Report, *Op. cit.*, pp. 52-57.

The specific assumption underlying the AGREE model is to require that each module makes an equal contribution to system mission success; in other words, each module has the same failure rate or mean life regardless of its specific function in the unit.

Using the approximation that $e^{-x} \approx 1 - x$, for small value of x, the allocated failure rate of the jth unit, $\hat{\lambda}_j$, is given in the AGREE report as

$$\hat{\lambda}_j = \frac{n_j[-\log R^*(T)]}{E_j\ t_j\ N} \tag{1-1}$$

where

$R^*(T)$ = reliability requirement for the system for T hours of system operation

n_j = the number of modules in the jth unit

E_j = the importance factor of the jth unit

$= \dfrac{\text{number of system failures due to jth unit failures}}{\text{number of jth unit failures}}$

t_j = the number of hours that the jth unit will be required to operate in T system hours

N = total number of modules in the system

The allocated reliability for the jth unit for t_j operating hours, $\hat{R}(t_j)$, is given by

$$\hat{R}(t_j) = 1 - \frac{1 - [R^*(T)]^{n_j/N}}{E_j} \tag{1-2}$$

The AGREE report has also pointed out some difficulties with using the module-count method of reliability allocation, stating that the method of reliability allocation discriminates against manufacturers developing a successor equipment to one which has already had its tube count reduced as compared to other equipment performing the same functions.[9]

ARINC Research Corporation has developed the basic allocation methods for serial, modified serial, and simple parallel systems.[10] The methods are basically a modification and extension of that presented in the AGREE report, and many of the AGREE recommendations for further study have been followed and solutions obtained. The major modification includes the adoption of the active element groups (AEG) concept of system definition, where the AEG is the smallest practical functional block which could be economically considered, and which would not be specifically related to existing configurations.

An active element is defined as a device which controls or converts energy; and AEG consists of one active element and a number of passive elements normally required to perform a specific system function. Because of the choice of the AEG as the basis of allocation, the AGREE requirement that each module makes an equal contribution to mission success is not necessary for this model.

---

[9] *Ibid.*, p. 56.

[10] ARINC Research Corporation, *The Allocation of System Reliability*, Technical Documentary Report, No. ASD-TDR-62-20, Vol. I (Washington, D. C.: ARINC Research Corporation, June, 1962).

Other major modifications include in the allocation model the distinction between functions, allowance for active element type, and the inclusion of the design adequacy concept of the system. The ARINC allocation models cover the case of a single redundant and bimodal system which represent a significant improvement over those presented in the AGREE report.

Several papers have recently appeared in the publications which present the allocation models based primarily on cost considerations.[11] The general approach used may be summarized as follows: A system reliability requirement of R^* exists for a system composed of n serial units. This reliability requirement is to be apportioned over n constituent units such that the allocated unit reliability of the jth unit, R_j, must satisfy the following relation

$$R^* = R_1 \cdot R_2 \cdot R_3 \ldots R_n$$

The cost of achieving a unit reliability of R_j, for $j = 1,2,\ldots,n$, is assumed to be $C(R_j)$. The allocation problem is then to find a set of minimizing values for R_j for the total cost function C given by

---

[11] A. M. Breipohl, "A Unique Allocation of Required Component Reliability," *Proceedings of the Seventh National Symposium on Reliability and Quality Control* (January, 1961), pp. 189-202; A. J. Truelove, "Mathematical Models for Optimizing Strategic Reliability and for Minimizing Costs," *Proceedings of the Sixth Joint Military Industry Guided Missile Reliability Symposium*, Vol. 2 (February, 1960), pp. 87-108; Fox Barbara, "Total Annual Cost, A Reliability Criterion," *Proceedings of the Tenth National Symposium on Reliability and Quality Control* (January, 1964), pp. 266-273.

$$C = C(R_1) + C(R_2) + \ldots + C(R_n) \tag{1-3}$$

subject to the reliability constraint

$$R^* = \prod_{j=1}^{n} R_j \tag{1-4}$$

The primary problem associated with the above approach is to introduce the relevant cost functions which will describe reality of the system's response to its cost measures, and yet be mathematically amenable. The cost functions proposed by Breipohl, Truelove, and Barbara are respectively as follows:

$$\left.\begin{aligned} C(R_j) &= \frac{A_j}{1 - R_j} e^{-B_j(1-R_j)} \\ C(R_j) &= \frac{A_j}{(1 - R_j)K} \\ C(\lambda_j) &= \frac{K_j}{\lambda_j} \end{aligned}\right\} \tag{1-5}$$

where A_j, B_j and K_j are the constants, and λ_j is the allowable unit failure rate of the jth unit within a system.

Carhart and Herd[12] have proposed a simple cost model to describe

---

[12] R. R. Carhart, and G. R. Herd, "A Simple Cost Model for Optimizing Reliability," *Reliability of Military Electronic Equipment, Appendix F, AGREE Report* (June, 1957).

a means of allocating expenditures between development and operating requirements by selecting that level of reliability which will minimize the total cost. The total cost per mission, C_M, is given in terms of the unit reliability as

$$C_M = C_B + \frac{C_F}{R} + C_b + d\left(\frac{-nt}{\log R}\right)^a \qquad (1\text{-}6)$$

where

C_B = basic cost of investment

C_F = variable cost of operation when equipment reliability were unity

C_b = cost of administration and development facilities

n = number of modules, a measure of complexity

T = module mean time to failure

R = $c^{-nt/T}$ = mission reliability of t hours

Albert and Proschan[13] have developed a mathematical model for optimal allocations of effort among constituent subsystems during the course of a development program in which system reliability is to be increased to a desired level at minimum total expenditure of effort. More specifically, the optimization problem is to minimize the total

---

[13] A. Albert, *A Measure of the Effort Required to Increase Reliability*, Technical Report No. 43, Applied Mathematics and Statistics Laboratory, Stanford University (November, 1958); A. Albert, and F. Proschan, *Increased Reliability with Minimum Effort*, Technical Report No. 50, Applied Mathematics and Statistics Laboratory, Stanford University (October, 1959).

effort function

$$G(R,R^*) = \sum_{j=1}^{n} G_j(R_j,R_j^*) \tag{1-7}$$

subject to the relation

$$\prod_{j=1}^{n} R_j^* = R^* \tag{1-8}$$

where $G_j(R_j,R_j^*)$ represent the effort function which measures the amount of effort needed to increase the reliability of the jth unit from R_j to R_j^*, and R^* is the desired system reliability requirement. In order to obtain an algorithm for the solution to the above problem, Albert and Proschan place the following restrictions on the effort function:

(1) $G(x,y) \geq 0$

(2) $G(x,y)$ is non-decreasing in y, and non-increasing in x

(3) If $x \leq y \leq z$, then $G(x,y) + G(y,z) = G(x,z)$

(4) $G(0,x)$ has a continuous derivative $h(x)$ such that $x \cdot h(x)$ is strictly increasing in $(0 < x < 1)$

With the above restrictions, the unique solution to the problem, after ordering the R_j's in non-decreasing order, is given by

$$R_j^* = \begin{cases} R_o^* & \text{if } j \leq K_o \\ R_j & \text{if } j > K_o \end{cases} \tag{1-9}$$

where K_o and R_o^* are determined respectively by the following relations:

K_o = maximum value of j such that

$$R_j < \left[\frac{R^*}{\prod_{K=j+1}^{n+1} R_K} \right] \tag{1-10}$$

where $R_{n+1} = 1$ by definition, and

$$R_o^* = \left[\frac{R^*}{\prod_{j=K_o+1}^{n+1} R_j} \right]^{\frac{1}{K_o}} \tag{1-11}$$

A number of investigators have been concerned with developing algorithms for obtaining the optimum redundancy to be used within a given system as a means of improving system reliability goal. The problem associated with the optimization of redundancy may be formulated as follows: Consider a system consisting of n serial subsystems with the reliabilities $R_1, R_2, \ldots, R_n$, which function satisfactorily if and only if each subsystem functions. It is assumed that at the jth stage a redundancy cost C_j units of money. The problem is then to find the method of obtaining optimum redundancy which will result in the maximum system reliability while not exceeding the total cost of given amount C. This formulation leads to the maximization of system reliability R(M) given by

$$R(M) = \prod_{j=1}^{n} [1 - (1-R_j)^{m_j}] \qquad (1\text{-}12)$$

with respect to n-component vector $M = (m_1, m_2, \ldots, m_n)$, denoting the level of redundancy to be used in each stage subject to the linear constraint

$$\sum_{j=1}^{n} C_j \, m_n \leq C \qquad (1\text{-}13)$$

This problem has been treated by Moskowitz and McLean,[14] using the Lagrange multiplier method. The same problem has been considered by Mine,[15] assuming that redundancy vector $M = (m_1, m_2, \ldots, m_n)$ are continuous variables. Mine's algorithm for finding the optimum redundancy vector M can be written as

$$m_j = \frac{1}{\log(1 - R_j)} \cdot \log \left[\frac{\lambda C_j}{\lambda C_j + \log(1/(1-R_j))} \right] \qquad (1\text{-}14)$$

where λ is the Lagrange multiplier satisfying the relation

$$\sum C_j \, m_j = C$$

---

[14] F. Moskowitz, and J. B. McLean, "Some Reliability Aspects of System Design," *IRE Transactions on Reliability and Quality Control*, PGRQC-8 (September, 1956), pp. 50-59.

[15] H. Mine, "Reliability of Physical System," *IRE Transactions PGIT, IT-5, Special Supplement* (1959).

If the subsystem reliabilities R_j are very close to 1, Mine shows that the value of the Lagrange multiplier is approximately given by the relation

$$\lambda = \exp(C) - \left[\prod_{j=1}^{n} \left[\frac{C_j}{\log \frac{1}{1 - R_j}} \right]^{\frac{1}{\log (1 - R_j)}} \right]^{\frac{1}{\sum_j \log (1 - R_j)} - 1} \qquad (1\text{-}15)$$

However, an exact solution for this case has been obtained by Kettelle, using the stepwise dynamic programming approach.[16] Basically, the Kettelle algorithm generates undominated redundancy allocations for successively larger subsystems from undominated allocations for small subsystems. The procedure requires that one first obtains a family of undominated redundancy allocations for the subsystems consisting of stage 1 and 2 alone, and then obtains the undominated allocations for the subsystems consisting of the stage 1, 2, 3, and 4, by using the undominated allocations of stages 1 and 2 already obtained in the previous step, and the undominated allocations for the combined stage 3 and 4 in the same way. In a similar fashion, combine stages 5, 6, 7, 8, etc., until all stages have been combined, resulting in a sequence of optimal solution to the problem.

Everett has also treated the redundancy allocation problem

---

[16] J. D. Kettelle, "Lease Cost Allocation of Reliability Investment," *Journal of the Operations Research Society of America*, Vol. 7, (1959), pp. 581-588.

using the generalized Lagrange multiplier method.[17] Each of the above allocation methods for optimum redundancy are applicable for the case of active parallel redundancy, and that the cost functions considered are of the linear form.

Flehinger[18] has discussed the optimal redundancy allocation problem for certain relay systems having sensing and switching devices. Geisler and Karr,[19] and Gourary[20] have treated the standby redundancy problems in which they minimize the expected value of shortage weighted by the essentiality of the item short subject to a single constraint of cost or weight.

Black and Proschan have made extensive studies of the standby redundancy allocation problems. In their several papers on this subject, they have been concerned with the determination of the number of spare parts of each type of components of a system which is required to give a specified assurance of continued operation during a given

---

[17] H. Everett, "Generalized Lagrange Multiplier Methods for Solving Problems of Optimum Allocation of Resources," *Journal of the Operations Research Society of America*, Vol. 2 (1963), pp. 399-415.

[18] B. J. Flehinger, "Reliability Improvement Through Redundancy at Various System Levels," *IBM Journal of Research and Development*, Vol. 2 (April, 1958), pp. 148-158.

[19] M. A. Geisler, and H. W. Karr, "The Design of Military Supply Tables for Spare Parts," *Journal of the Operations Research Society of America*, Vol. 4 (1956), pp. 431-442.

[20] M. H. Gourary, "A Simple Rule for the Consolidation of Allowable Lists," *Naval Research Logistics Quarterly*, Vol. 5, No. 5 (1958), pp. 1-15.

period of system operation at minimum cost.[21] More specifically, Black and Proschan have treated a system consisting of K serial components whose failure distribution for the jth subsystem is $F_j(t)$ with all components independent. It is further assumed that during the period considered only the spares initially provided may be used to replace components that have failed.

The problem is then to select a set of non-negative number n_j of redundant units to maximize the probability of survival until time t_o, that is, to maximize the quantity given by

$$F(N) = \prod_{j=1}^{K} F_j(n_j) \tag{1-16}$$

subject to linear constraint

$$C(N) = \sum_{j=1}^{K} n_j C_j \leq C_o \quad \text{and} \quad n_j \geq 0 \tag{1-17}$$

$$\text{for } j = 1,2,\ldots,K$$

where

$F_j(n_j)$ = probability that n_j or less failures of jth subsystem occur during $(0,t_o)$, $j = 1,2,\ldots,K$

[21] G. Black, and F. Proschan, "On Optimal Redundancy," *Journal of the Operations Research Society of America*, Vol. 7 (1959), pp. 581-588; G. Black, and F. Proschan, "Spare Parts Kits at Minimum Cost," *Proceedings of the Fifth National Symposium on Reliability and Quality Control* (January, 1959), pp. 281-289; F. Proschan, *Minimum Cost for Spares When Component Failure is not Necessary Exponential*, Report EDL-M 185, Electronic Defense Laboratory, Mountain View, California (1959).

N = $(n_1, n_2, \ldots, n_K)$, the vector specifying the number of spares of each subsystem

$F(N)$ = probability that no shortage is experienced for any of K subsystems during $(0, t_o)$, given an initial kit of composition N

$C(N)$ = total cost of a spare-parts kit composed of n_j units of jth subsystem for $j = 1, 2, \ldots, K$

C_o = allowable total cost of the spare parts

The solution algorithm for the case of one linear constraint is obtained by Proschan, assuming the exponential failure distribution.

The solutions of this standby redundancy allocation problems have since been extended to a wide class of failure distributions other than the exponential. In particular, Proschan[22] has shown that the method of solution is also valid for the case where underlying failure distributions are polya frequency functions of Type 2, that is, the f(t) have the property such that the relation

$$\begin{vmatrix} f(x_1 - y_1) & f(x_1 - y_2) \\ f(x_2 - y_1) & f(x_2 - y_2) \end{vmatrix} \geq 0 \qquad (1\text{-}18)$$

is satisfied for $x_1 < x_2$ and $y_1 < y_2$.

A number of papers employing the dynamic programming approach to the problem of optimum allocations of redundancy in the serial system

---

[22]F. Proschan, "Optimal System Supply," *Naval Research Logistics Quarterly*, Vol. 7, No. 4 (December, 1961), pp. 609-646.

have appeared in recent publications: Bellman and Dreyfus[23] have treated the problem of maximizing system reliability concurrently with two linear constraints of cost and weight. Liittschwagner has also considered the dynamic programming formulation of finding maximum system reliability concurrent with the minimization of weight and power for the case of standby redundancy.[24] Sasaki has also treated the redundancy allocation problem when there are n dimensional constraints such as cost, weight, and volume.[25]

The dynamic programming methods employed by the above authors may be summarized as follows: it is assumed that one component of the jth type must be used at each stage. Stage j consists of $n_j + 1$ units in parallel, each of which has independent probability q_j of failure. Practical constraints of cost, weight or size exist in a given system. Let the cost of a single component of the jth stage be c_j and w_j be its weight. The problem is then to find the number of redundant units which will result in maximum system reliability consistent with the given linear constraints. This formulation leads to the maximization of system reliability R(N), where $N = (n_1, n_2, \ldots, n_K)$, given by

---

[23] R. Bellman, and S. Dreyfus, "Dynamic Programming and the Reliability of Multicomponent Devices," *Journal of the Operations Research Society of America*, Vol. 6 (1958), pp. 200-206.

[24] J. M. Liittschwagner, "Dynamic Programming in the Solution of Multi-Stage Reliability Problem," *The Journal of Industrial Engineering*, Vol. XV, No. 4 (July-August, 1964), pp. 168-175.

[25] M. Sasaki, "A Simplified Method of Obtaining Highest System Reliability," *Proceedings of the Eighth National Symposium on Reliability and Quality Control* (January, 1962), pp. 489-502.

$$R(N) = \prod_{j=1}^{K} (1 - q_j^{\,n_j+1}) \tag{1-19}$$

subject to the linear constraints

$$\sum_{j=1}^{K} c_j n_j \le C$$

$$\sum_{j=1}^{K} W_j n_j \le W \quad \text{and} \quad n_j = 0,1,2,\ldots \qquad j = 1,2,\ldots,K. \tag{1-20}$$

In general, the above problem is a typical example in the theory of non-linear programming where the decision variable at each stage must be a non-negative integer, and a general solution in closed form cannot be given. The solutions derived by the authors are given in terms of a computational algorithm only.

More recently, Proschan and Bray[26] have extended the Ketteles procedure which generates a family of undominated redundancy allocations when more than one linear cost constraints are present. Barlow, Hunter, and Proschan[27] have discussed the problem of optimal redundancy alloca-

---

[26] F. Proschan, and T. A. Bray, "Optimum Redundancy Under Multiple Constraints," *Boeing Scientific Research Laboratories Document DI-82-9253* (1963).

[27] R. E. Barlow, L. C. Hunter, and F. Proschan, "Optimum Redundancy When Components are Subject to Two Kinds of Failures," *Journal of the Society for Industrial and Applied Mathematics*, Vol. 11, No. 1 (March, 1963), pp. 64-73.

tions when components are subject to two kinds of failures such as an open circuit failure and short circuit failure in the network of relays.

Objectives

Perhaps the most important reason for undertaking the present study is that the further need exists for developing a comprehensive reliability allocation model which is generally applicable, is economically feasible, and yields realistic reliability requirements. Each of the above-cited approaches to the reliability allocation problem has emphasized only certain aspects of the overall system reliability allocation problem. The basic allocation model, for example, has tended to focus only on the unit capability based on the historical failure information while neglecting the other important system characteristics such as the development cost and physical constraints which are always present in any allocation process.

On the other hand, a class of redundancy allocation problems studied constitutes a problem of allocating redundancy in order to achieve or improve system reliability goal within a system, and these approaches are not truly a problem of allocating system reliability requirements since no considerations are given to the unit state-of-the-art, its functional importance, and environmental factors which will affect the overall system reliability. Furthermore, a class of reliability allocation problems as viewed from the overall system optimization approach has remained largely untouched and much needs to be explored. This points up the need for a comprehensive model of the reliability allocation which takes into account a large number of interdependent and interacting factors contained in a given reliability allocation problem.

The general objective of this research is to develop a methodology for assigning a specific system reliability requirement to the various components of a system when it is desired to achieve a specified level of reliability goal consistent with a set of system constraints. Specific areas of investigation include the following:

1. The development of a procedure for determining the optimum level of reliability requirements for each of the major functional units which comprise a complex system.

2. The development of a detailed allocation procedure for determining the optimum design policy and the quantity of redundant units to be used within a system which will minimize the cost of achieving a specified level of reliability consistent with the given system constraints.

3. The development of a procedure for allocating system reliability requirement into subsystem levels based on the concept of the total system costs.

4. The development of appropriate computational procedures for the reliability allocations and associated redundancy problems.

5. The evaluation of reliability improvement obtainable through repairable redundancy at the weak system elements.

CHAPTER II

FUNCTIONAL ALLOCATION OF SYSTEM RELIABILITY

Basic Considerations

The reliability requirement for a new system or equipment is usually established during the early development phases of a system design. This reliability must, in turn, be apportioned or allocated into the numerous components, units, and subsystems that make up a system in order to determine the reliability goals that these units must achieve if the overall system requirement is to be met. These allocations are accomplished through the use of an appropriate allocation procedure.

For any system composed of n units, the reliability allocation process may be symbolically expressed as the following functional inequality:

$$R^{*}(t) \leq \phi\left(R_1(t), R_2(t), \ldots, R_n(t)\right) \tag{2-1}$$

where

$R^{*}(t)$ = the system reliability requirement for t system hours of operation

$R_j(t)$ = the allocated reliability for the jth unit for a time period of t hours, and

ϕ denotes the functional relationship between system and unit reliabilities.

For a simple serial system where failure of any constituent unit will result in failure of the whole system, and failure probability of each unit in the system is assumed to be independent, the above relation reduces to

$$R^{*}(t) \leq R_1(t) \cdot R_2(t) \cdots R_n(t) \qquad (2\text{-}2)$$

Thus the overall reliability of the system can be expressed as the product of the reliabilities of the constituent units.

Theoretically, there exist a large number of feasible solutions to the above reliability equation, assuming no restrictions imposed on the allocation process. A very simple solution would involve apportioning the system reliability requirement into n equal parts so that allocated reliability would be identical for all units in the system. Obviously, this is not a realistic approach because complex units would have the same reliability requirements as relatively simple units. Therefore, some restrictions must be imposed on the solution of this equation if consistent and realistic reliability allocations among units are to be desired.

In order to obtain realistic solutions, the allocation process basically involves the identification and appropriate quantification of these factors which influence the unit reliability, and establishes a procedure capable of producing a unique or limited number of solutions consistent with the appropriate system constraints.

A class of allocation methods currently available gives by itself no assurance or guarantee that the reliabilities so assigned to the units will materialize in service operation of the unit or system. The reliability allocations arrived at by the current method often need to be compromised through the trade-off analysis between reliability, other performance requirements, weight, space, development time, and more frequently monetary limitations.

Furthermore the current allocations methods are not suited for treating more than two redundant units because the complexity of the allocation model is greatly increased for more than one redundant configurations. Both of these reasons point out the need for a comprehensive model of the reliability allocation which takes account of these factors along with other important system characteristics.

The general criteria employed in the development of such a comprehensive model may be outlined as follows:

1. The method must be generally applicable to a variety of system configurations, and all classes of system elements--electronic, electro-mechanical, mechanical, hydraulic, chemical, etc., and furthermore it must be applicable at the various phases of a system development process.

2. The method must be based on the available input data in the early phases of system design. Since system reliability requirements are normally allocated into the subsystems during the early phases of development program, the definition and quantification of allocation factors incorporated in the model are consequently dependent upon the available design information during this phase of the development program.

3. The method must provide economically feasible and realistic allocations in terms of development funding, time schedule, state-of-the-art or other design requirements.

Reliability Functions

The system reliability requirement constitutes the basic parameter in a reliability allocation model. This requirement is usually established by either the prospective user or proposed by the supplier of the equipment or system in consideration of such factors as mission requirements, operational readiness, maintainability, safety, system cost, development schedule, and other factors of intrinsic importance. The requirement is normally specified in any appropriate measure such as system failure rate, mean life or reliability of the system for a fixed duration of system operation.

The most widely accepted definition of reliability is given as the probability that a system will perform satisfactorily for at least a given period of time when used under stated conditions.[1] When applied to a specific one-shot type system such as missiles or satellites, mission reliability is frequently defined as the probability that a system will operate in the mode for which it was designed for the duration of a mission, given that it was operating in this mode at the beginning of the mission.[2] Thus, mission reliability is a measure of probability of survival of the system for the period of time required to complete the

---

[1] William H. Von Alven, *Reliability Engineering*, (Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1964), p. 6.

[2] *Ibid.*

specified mission.

Whenever the reliability definition is established to fit a particular system or equipment, it is then necessary to base the measure of probability on the precise definition of "success" or "satisfactory performance" of system function, and then choose the time period or operating cycle over which such satisfactory system operation is desired, and finally specify the environmental or use conditions under which the system is expected to operate.

The reliability function for a given system represents this same probability of successful system operation expressed as a function of the time period of interest. The general properties that the reliability function, $R(t)$, is assumed to satisfy are:

$$R(0) = 1 \quad \text{and} \quad R(\infty) = 0 \tag{2-3}$$

and $R(t)$ is a monotone nonincreasing function in the time interval $(0,\infty)$.

For a given system, the reliability function $R(t)$ is related to the failure density function $f(t)$ by

$$R(t) = \int_t^\infty f(t)dt = 1 - \int_0^t f(t)dt \tag{2-4}$$

Differentiating both sides of Equation (2-4) gives

$$R'(t) = - f(t), \tag{2-5}$$

assuming that the derivative exists. By definition, $f(t)dt$ gives the probability that the system starting at time $t = 0$ will fail in the time interval $(t, t + dt)$. Then, the conditional probability that the system will fail in $(t, t + \Delta t)$, given that it is operating at time t is given by the quantity

$$\frac{R(t) - R(t + \Delta t)}{R(t)} \qquad (2\text{-}6)$$

The hazard rate or instantaneous failure rate is defined to be

$$h(t) = \lim_{\Delta t \to 0} \frac{R(t) - R(t + \Delta t)}{\Delta t \cdot R(t)} = \frac{f(t)}{R(t)} \qquad (2\text{-}7)$$

or

$$h(t) = - \frac{d[\log R(t)]}{dt} \qquad (2\text{-}8)$$

From the relationship in (2-7) and using (2-5), $R(t)$ can be written in terms of the hazard rate as

$$R(t) = \exp\left(-\int_0^t h(x)dx\right) \qquad (2\text{-}9)$$

which holds for any form of hazard rate function, and is thus called the general reliability formula. It is seen from Equation (2-9) that as $t \to \infty$, $R(\infty) = 0$, which agrees with the assumption given in (2-3). The Kth moment of the failure density is usually denoted by u_K' and is defined as

$$u_K' = \int_0^{\infty} t^K f(t)dt$$

or, upon integration by parts, it follows that

$$u_K' = K \int_0^{\infty} t^{K-1} \cdot R(t)dt \qquad (2\text{-}10)$$

Some reliability functions which have been frequently used in the reliability studies are as follows.[3]

(1) Exponential distribution

$$R(t) = e^{-\lambda t}, \quad \lambda > 0$$

$$h(t) = \lambda$$

$$u_K' = \frac{K!}{\lambda^K}$$

(2) Gamma distribution

$$R(t) = \frac{1}{\Gamma(m+1)} \int_{\lambda t}^{\infty} u^m e^{-u} du$$

where $m > -1$, and $\lambda > 0$

$$h(t) = \frac{\lambda^{m+1} t^m}{(m+1)} \cdot \frac{e^{-\lambda t}}{R(t)}$$

[3]R. E. Barlow, and F. Proschan, *Mathematical Theory of Reliability* (New York: John Wiley & Sons, Inc., 1965), p. 13.

$$u'_K = \frac{\Gamma(m + K + 1)}{\Gamma(m + 1)} \cdot \frac{1}{\lambda^K}$$

(3) Weibull distribution

$$R(t) = \exp(-t\ a/\theta),\ a,\ \theta > o$$

$$h(t) = \frac{at^{a-1}}{\theta}$$

$$u'_K = \Gamma\left(\frac{K}{a} + 1\right)\ \theta^{\frac{K}{a}}$$

(4) Truncated normal distribution

$$f(t) = \frac{a\phi\{a(t-t_o)\}}{\Phi(at_o)} \qquad a > o,\ -\infty < t_o < \infty$$

$$\text{where} \quad \phi(x) = \frac{1}{\sqrt{2\pi}}\ e^{\frac{-x^2}{2}}$$

$$\text{and} \quad \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-t^2}{2}}\ dt$$

$$h(t) = \frac{a\phi\{a(t-t_o)\}}{\Phi\{-a(t-t_o)\}}$$

$$u'_1 = t_o + \frac{1}{a} \cdot \frac{\phi(a\ t_o)}{\Phi(a\ t_o)}$$

$$u'_2 = t_o^2 + \frac{1}{a^2} + \frac{t_o}{a} \cdot \frac{\phi(a\ t_o)}{\Phi(a\ t_o)}$$

For a system composed of n units in series where the failure density function of each unit within the system is assumed to be independent, and furthermore the failure of any one unit causes a failure of the system, the system reliability which gives the probability that every unit will operate without failure for at least time t, can be obtained by the product of individual unit reliabilities. Denoting by $R_j(t)$ the reliability function of the jth unit and by $h_j(t)$ the hazard rate function for $j = 1,2,3,\ldots,n$, then the serial-system reliability function has the form

$$R(t) = \prod_{j=1}^{n} R_j(t)$$

But, it follows from Equation (2-9) that $R_j(t)$ can be written as

$$R_j(t) = e^{-\int_o^t h_j(x)dx}$$

Hence

$$R(t) = e^{-\left[\sum_{j=1}^{n} H_j(t)\right]} \tag{2-11}$$

where $H_j(t) = \int_o^t h_j(x)\,dx$; for $j = 1,2,\ldots,n$.
Equation (2-11) points out the fact that the system hazard rate is the sum of the unit hazard rates associated with (0,t) under the assumption of independence of unit failure, and that this relationship holds regardless of the form of the unit failure distributions.

Thus, knowledge of the hazard rates of constituent units is often useful in the prediction of a system reliability. Since the system hazard rate is the sum of unit hazard rates for any form of unit reliability function, it is logical to add unit hazard rates and use it as an approximation of the system hazard rate in a given prediction, assuming the independence of each unit failure.

Exponential Distribution as a Failure Law of Complex Systems

The usual practice in reliability analyses on any equipment or system is first to assume an appropriate failure distribution which will describe mathematically the length of life of a system under consideration. The modes of possible failure for the system will also significantly affect the analytic form of the failure distribution.

Although there is no clear distinction between these modes of failure in practice, three dominant modes of failure for a given system are generally recognized throughout the system life cycle. They are called "early failure," "chance failure," and "wear-out failure." Figure 2 from Moskowitz shows a typical failure rate of a newly-developed system expressed as a function of time.[4]

When a newly-developed system is first placed into operation, the system may initially exhibit a relatively high failure rate if it contains some design defects or poor workmanship. As these defectives or weak units fail one by one, and are replaced with new ones, the failure rate will decrease rapidly during the so-called "debugging"

---

[4]Fred Moskowitz, "The Statistical Analysis of Redundant Systems," *IRE International Convention Record*, Part 6 (March, 1960), pp. 78-89.

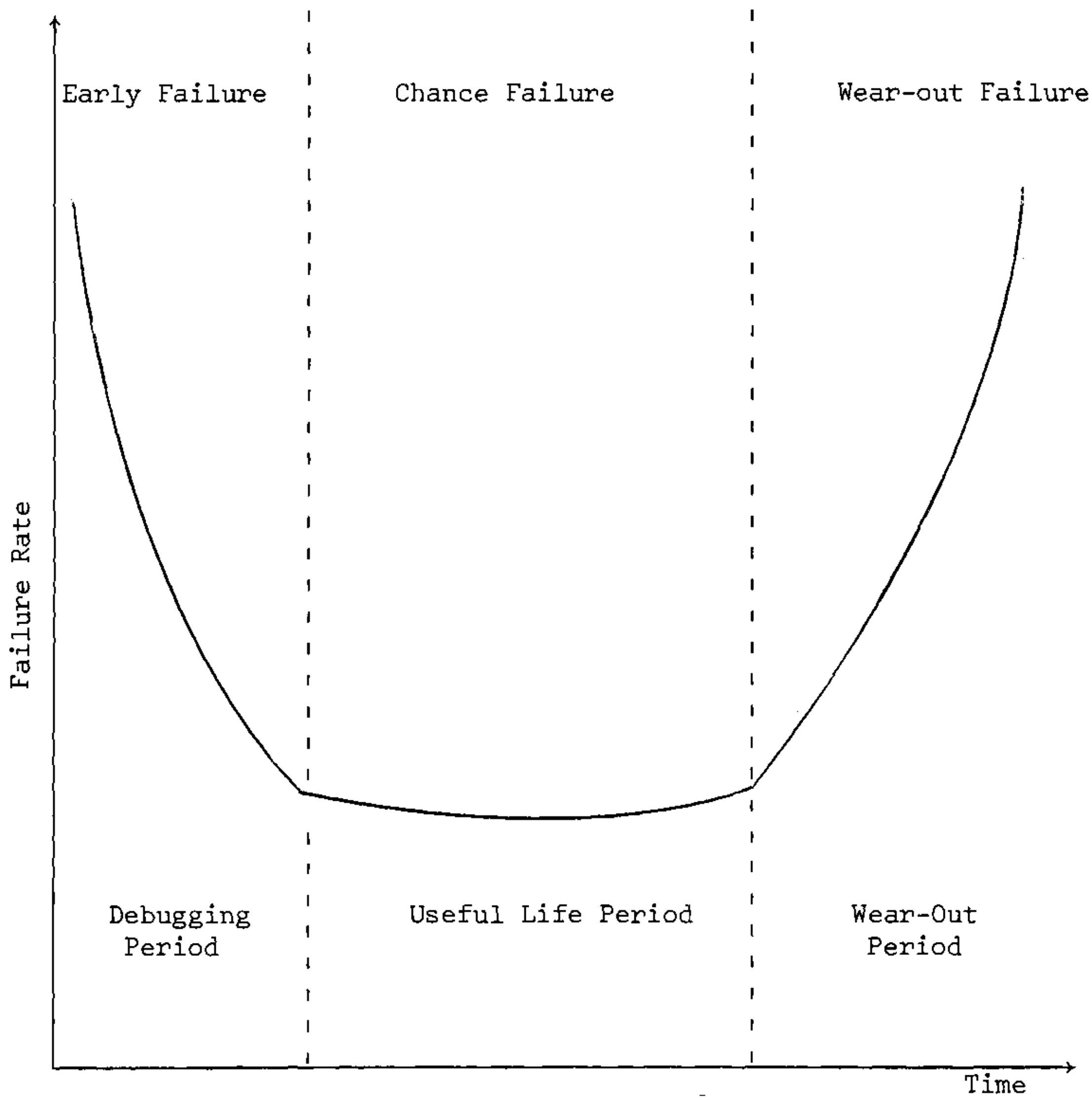


Figure 2. Typical Failure Rate of Newly-Developed Systems.

period, and gradually stabilize to an approximately constant level. During this constant failure rate period the system is expected to operate to the full extent for its intended purpose. This period of the system life cycle is often called useful life period. As the system approaches to the end of its useful life cycle, failure rate begins to rise because of gradual deterioration and wear-out of its component.

In most cases, early failure of the system results from improper design, poor manufacturing and quality control techniques during the design and production processes, and even improper use (environmental) conditions. These failures can be eliminated by employing a proper debugging technique in which the system is placed in operation for a number of hours under specified conditions simulating the actual use environments, and any defect found is replaced by good components or corrected before it is released for actual service.

On the other hand, wear-out failures are caused by deterioration of components due to aging. These failures often occur in a system where maintenance is not performed properly or maintenance is not possible at all, such as the case of the one-shot system. No matter how reliably the system is built initially, the system will begin to deteriorate soon or later unless it is properly maintained. Wear-out failure in most cases can be prevented through proper methods of component preventive maintenance or replacement so that components do not get a chance to fail because of wear-out and deterioration.

Chance failure, sometimes called catastrophic failure, still occurs at random within the useful life period of a system life cycle even though all efforts have been made to eliminate design defects and weak components, and before wear-out becomes predominant. These failures are characterized by a sudden breakdown of the system, without preceding noticeable deterioration symptoms, caused by instant stress accumulations beyond the design strength of the system.

When a system is placed into operation under actual environmental conditions in which stress accumulations that may cause component failures occur at essentially random, then the component failures themselves also will occur at random. As long as the system is properly maintained by replacing failed components with new ones, approximately the same number of failures will occur in periods of equal length.

The exponential distribution is widely used in reliability studies on complex systems. This distribution is a direct consequence of the assumptions that the probability of system failure in a given time interval is directly proportional to the length of the interval, and is independent of the age or past history of the system. The exponential density function of time to failure from this assumption has the form

$$f(t) = \lambda e^{-\lambda t}$$

The failure rate is equal to λ, and is inversely proportional to mean time to failure.

The property of exponential distribution implies two significant failure characteristics; first, individual failures occur at random, and secondly, the failure rate is constant, which in turn implies that wear-out or deterioration is not a failure cause in a given system. It can be proved that if the life length t of a system has the exponential distribution, then past use history of the system does not affect its future life length of the system.[5] In other words, if a system has not failed up to a time t_o, the probability distribution of its future life length $t-t_o$ is the same as if the system were new and had just been placed into service at time t_o. The only distribution satisfying this condition is the exponential, and for this reason, the exponential distribution is often called the distribution with complete lack of memory.[6]

A number of mathematical arguments supported by some empirical evidence have appeared to support the adequacy of the exponential distribution as a failure distribution for a complex system. The use of exponential distribution as a failure law of a complex equipment can be traced back to the earliest work of Khintchine[7] and C. Palm,[8] in which they have justified the exponential distribution as an input distribution of time between machine breakdown in the automatic machine maintenance problem.

---

[5] William Feller, *An Introduction to Probability Theory and its Applications* (New York: John Wiley & Sons, Inc., 1953), p. 413.

[6] *Ibid.*, p. 412.

[7] A. Ya. Khintchine, "Mathematisches Uber die Erwartung von einen Offentlicher Schalter," *Matem. Sbornik* (1932).

[8] Docent C. Palm, "Arbetskraftens Fordelning vid betjaning av automatskiner," *Industritidningen Norden*, Vol. 75 (1947), pp.75-80, 90-94.

Davis has studied the failure data for a wide variety of electronic components, and noted that the results of several goodness-of-fit tests for various competing failure distributions tend to substantiate the hypothesis of the exponential failure distribution.[9] For this reason, many references have been made to the David paper to support the assumption of the exponential distribution in the reliability work.

Epstein and Sobel are generally credited with thorough exploration of the family of exponential distributions for use in the life testing research. Their important studies in the field of life testing have advanced the widespread assumption of exponential distribution in the reliability work.[10]

It is also noteworthy that many test programs prescribed by the Department of Defense for contractors are based on the assumption of exponential failure distribution.[11]

Except for a complex system containing a large number of components with component renewal policy, many other failure distributions have been proposed on the basis of empirical studies and mathematical convenience for different applications. Typical failure distributions are the Weibull, the normal, and the log normal. All of these failure

---

[9] D. J. David, "Analysis of Some Failure Data," *Journal of the American Statistical Association,* Vol. 47, No. 258 (1952), pp. 113-150.

[10] B. Epstein, and M. Sobel, "Life Testing," *Journal of the American Statistical Association,* Vol. 48, No. 263 (1953), pp. 486-502.

[11] See, for example, Department of Defense Handbook H 108, "Sampling Procedures and Tables for Life and Reliability Testing," April 29, 1960.

distributions, except the log-normal whose hazard rate increases at first and then decreases eventually, have in common increasing failure rates for some parameter values. Thus, this increasing hazard rate may frequently represent the reality of the underlying failure mechanism of the system. Since the hazard rate decreases with time in the log normal, this distribution has found more applications as a repair time distribution rather than a failure distribution.[12]

One of the basic reasons for a widespread use of the exponential distribution in the reliability studies lies in the fact that it lends itself a simple addition of failure rates in obtaining the failure rate of a complex system, and thus enables a data compilation in much simpler form than any other failure distribution. In fact, one author has suggested using the exponential as the touchstone of reliability.[13]

However, as Zelen and Dannemiller pointed out, the assumption of an exponential distribution may lead to serious errors if the underlying failure distribution is the Weibull distribution.[14] Birnfaum and Saunders have presented a statistical model for life lengths of structures under dynamic load in which they make it possible to express the probability distribution of life length in terms of the load given as a

---

[12] A. W. Green, and R. C. Horne, "Maintainability of Shipboard Electronic Systems," *ARINC Research Corporation Publication*, No. 118-4-228, Washington, D. C.

[13] C. M. Ryerson, "Reliability Testing Theory Based on the Poisson Distribution," *Proceedings of the Fourth National Symposium on Reliability and Quality Control* (1958), pp. 3-18.

[14] M. Zelen, and M. C. Dannemiller, "The Robustness of Life Testing Procedures Derived from the Exponential Distribution," *Technometrics*, Vol. 3, No. 1, pp. 29-49.

function of time and of deterioration occurring in time independently of loading. For the special case of a constant load or of periodic loading with constant amplitude, they suggest in certain situations the gamma distribution as an adequate distribution for life length.[15]

General Assumptions

The foregoing discussion has established the nature of the reliability allocation problem, and also pointed out the factors which should be considered in the allocation process. The reliability allocation of a large-scale system involves numerous complicated problems which are usually interrelated. The complexity of the allocation problem varies to the large extent with the specified mission requirements of the system, available knowledge of its feasibility and inherent limitations, and specific experience gained from the comparable systems previously developed.

The measures available to achieve very high reliability may conflict with other desirable system characteristics such as high performance, reduced weight, and low development cost. It is therefore apparent that a reasonable allocation model must recognize these interactions among various factors, and necessarily reflect the combined effects of these factors in the allocation process.

Because of the complexity involved in building such a realistic model, the system reliability allocation is carried out in this study through two distinct methods. The first method called the "functional allocation" method determines the preliminary apportionment of reliability requirements at the first primary level of breakdown of a system (e.g., major functional equipment groups) consistent with the requirement of the overall system reliability. The second method, called "detailed allocation" method, accomplishes the detailed, supplementary allocations within the functional equipment groups including the use of redundancy, and will be discussed in the following chapter.

Figure 3 displays the block diagram showing the primary and secondary levels of breakdown of a system configuration. Reliability allocations among the primary level (functional equipment groups) of breakdown of a system can be accomplished by the use of the functional allocation method. The detailed allocations of a system reliability requirements among the various units including redundancy within the functional equipment groups are made through the detailed allocation method.

The preliminary allocation of system reliability requirements at the first level of functional configurations of the system is necessary because each of the major functional equipment groups is often produced by a separate contractor, and frequently the prospective user, such as the Armed Forces for the weapons system, desires to specify a numerical requirement into each contract in order to ensure that the entire system, when assembled, would meet the overall system requirement. This would be the situation, for example, in the case of a missile system. The entire missile system must satisfy a certain reliability requirement which should then be apportioned into requirements for the propulsion contractor, requirements for the flight control contractor, etc. A still further breakdown of these major subsystems into major units and components might be made whenever necessary,

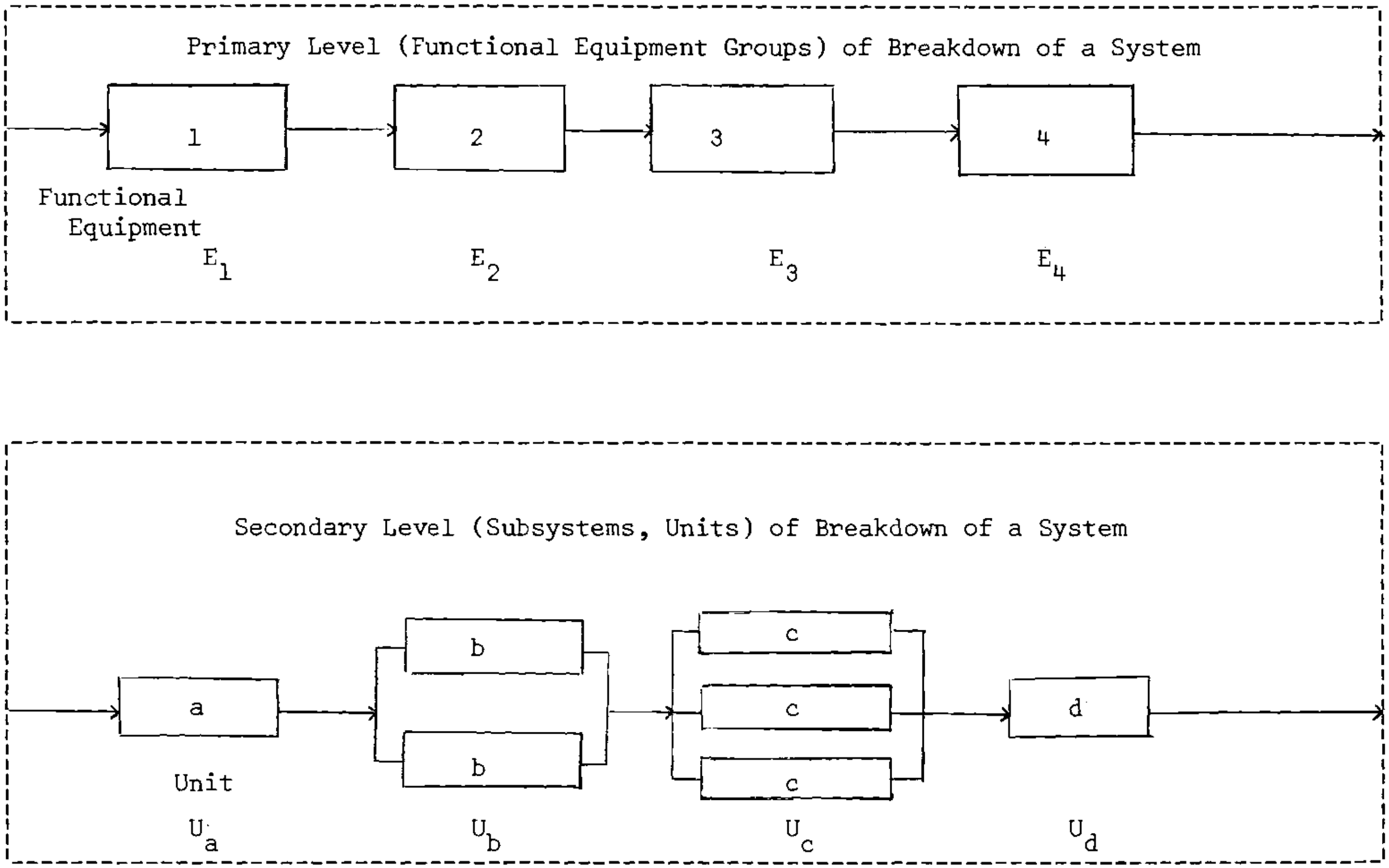


Figure 3. Block Diagram Showing the Primary and Secondary Levels of Breakdown of a System Configuration

and unit requirements could be established.

In the functional allocation, the system which includes redundancy is not considered because of the consideration that preliminary system design in the early phases of system development primarily determines the type and minimum set of the functional equipment required to perform the intended system functions, and thus, the use of redundancy is not a standard design practice at this stage of system design, since the redundancy techniques can be more effectively employed as a means of achieving allocated reliability requirements after the preliminary allocations are made and any deficiency area is pinpointed through the feasibility study.

The conceptual approach employed in the functional allocation method is that the multi-stage system is first divided into two or more functional equipment groups which can fail independently, and then the specified system reliability requirement or equivalent allowable failure rate is to be apportioned among those functional equipment groups by the use of the allocation index.

The allocation index used in this method represents a numerical measure which quantitatively characterizes those factors which significantly influence unit reliability. The allocation index which will be discussed in detail in the following section is used in the allocation method as the best engineering approximation of relative reliability of the allocation unit within a system.

The specific assumptions made in the allocation process are as follows:

A. System's reliability structure: The system under consideration sists of n functional equipment groups connected in series so that failure of any of the functional equipment in the system will (1) result in an entire system failure, or (2) not necessarily result in an entire system failure, i.e., the failure of a particular equipment will not lead to the failure of an entire system.

B. Independence of the failure probabilities of the functional equipment: The functional equipment groups within a given system are so chosen that the failure probability of each functional equipment is independent of all other functional equipment.

The above assumptions will simplify a mathematical model which describes the relationship between the functional equipment and the system reliability, and will allow to use the product formula for a system reliability function upon which the allocation method is based.

With regard to the first assumption, one example of the serial system is that of a complex missile consisting of several stages. In this missile system, there are groups of functional equipment, subsystems and units, some of which operate simultaneously while others operate sequentially, and all of which must function satisfactorily before the intended system mission can be successfully accomplished. Thus the first stage must operate satisfactorily, followed by successful stage separation and then by successful ignition and firing of the second stage engine, etc. However, the first stage itself is a serial system. It consists of several booster engines, guidance, and control subsystems so that each of these subsystems must function simultaneously and successfully for stage success. For all practical purposes, it is possible to group the major units of a given system into a limited number of the serial functional equipment groups, each of which is vital to the successful system operation.

On the other hand, it is sometimes unlikely that a system containing a large number of units will fail every time any one of the units fails. It is more likely that the failure of one or more units will cause some degree of system performance degradation rather than a catastrophic system failure. This assumption is particularly valid for the lower levels of breakdown of a system configuration. In this connection it is to be noted that it is more common for units to drift out of tolerance than it is for units to fail completely in a given system.

While the second assumption is mathematically convenient because the system reliability can be expressed as a product of independent reliabilities of the functional equipment rather than as a product of conditional probabilities, this assumption is still believed to be valid as far as the early design-stage reliability allocations among the major functional equipment are concerned. Generally, if a given system is broken down into a limited number of major functional equipment groups, then these functional equipment groups can be regarded as being independent with regard to their reliabilities. This assumption seems quite reasonable at the higher levels of breakdown of a system configuration because the functional equipment groups are frequently produced by the separate contractors, and the reliability numbers on which the allocation method is based is estimated from separate (independent) functional equipment life tests.

If the functional equipment groups within a system is known to be interdependent, these dependent groups may possibly be grouped into one allocation group, making the failure probability of this group independent of the state of other groups. For instance, if the failure of equipment E_2 in Figure 3 is known to be strongly dependent on the failure of equipment E_1 and is not likely to fail if E_1 does not malfunction, they can be grouped together, and the state-of-the-art of this group may then be adjusted to minimize the error in an allocation process. This approach is believed to be reasonable since the probability of successful operation of equipment E_2 is dependent on the successful operation of equipment E_1, and this probability is already contained in the adjusted probability of successful operation of the new group.

Generally, the effect of nonconformity to the second assumption in practice will result in allocation numbers which are higher than are actually necessary for obtaining the desired level of the system reliability requirements.[16] Therefore, the system designers are on the safe side of the error introduced by failure to conform to the assumption. If all functional equipment will operate satisfactorily as required by the allocation, the overall system reliability will exceed its actual reliability requirement.

Functional Allocation Model

Under the basic assumptions made above, the reliability of every complex system can be expressed as the product of the reliabili-

[16] Joan R. Rosenblatt, "On Prediction of System Performance from Information on Component Performance," *Proceedings of Western Joint Computer Conference* (February, 1957), pp. 85-94.

ties of all those units on whose satisfactory operation the system depends for its survival. For a system composed of n serial units in which some unit will operate all the time, while others will operate only for short time interval or only for a few cycles intermittently during a given system operating time t, the system reliability for a system mission time of t hours is given by

$$R(t) = R_1(t_1) \cdot R_2(t_2) \cdot \cdot \cdot R_n(t_n) = \prod_{j=1}^{n} R_j(t_j)$$

where $R_j(t_j)$ is the reliability of the jth unit which is required to operate t_j hours during t system operating hours. But, using the relationship obtained in (2-11), R(t) can be expressed in terms of unit hazard rate as

$$R(t) = e^{-\sum_{j=1}^{n} \int_{o}^{t_j} h_j(x_j)dx_j} \qquad (2\text{-}12)$$

where $h_j(t_j)$ is the hazard rate function of the jth unit at time t_j.

When the unit hazard rate function $h_j(t_j)$ is known to be constant for a given operating time t, the system's reliability for this time period can be immediately obtained from Equation (2-12) by simply summing the constant hazard rates. The operating time period t for which the reliability of the system is desired may be of any length as long as the hazard rate must not change during this time period contemplated.

For a given system this constant hazard rate implies that all the engineering efforts have been made in debugging and in preventing wear-out and deterioration, and consequently the system life will behave exponentially with regard to its failure under essentially the same environments, that is, the system will fail only by chance failure caused by sudden and unpredictable accumulations of various stresses beyond the design strength and application stress level of the units and components which form the system.

For those units whose hazard rates change with time, a reasonable approach would be to average the hazard rate over the unit operating period t_j and to replace the $h_j(x_j)$ in Equation (2-12) by this average hazard rate, provided that routine preventive maintenance is practiced in the system. The average hazard rate over the operating period t_j may be obtained by

$$\lambda_j = \frac{1}{t_j} \cdot \int_0^{t_j} h(x_j)dx_j \tag{2-13}$$

Throughout the operating period $(0,t_j)$, if the environmental stresses change from one level to another so that the jth unit in the system will exhibit a non-constant failure rate, but does show an approximately constant failure rate, say λ_j' for the operating time t_j' at the first stress level and another failure rate λ_j'' for the operating time t_j'' at the other stress level, then the equivalent average failure rate for the entire operating time $t_j = t_j' + t_j''$ may be obtained by

$$\lambda_j = \frac{\lambda_j' \, t_j' + \lambda_j'' \, t_j''}{t_j} \tag{2-14}$$

Under the assumption of exponential failure distribution in time interval $(0,t_j)$, $h_j(x_j)$ in Equation (2-13) can be replaced by the equivalent average failure rate λ_j and the system reliability function R(t) can now be written as

$$R(t) = e^{-\sum_{j=1}^{n} \lambda_j t_j} \tag{2-15}$$

Given a reliability requirement $R^*(t)$ for the system, the basic constraint in the allocation process that the allocated unit reliabilities, when recombined, must yield the overall requirement gives the following relationship

$$R^*(t) = e^{-\lambda^* t} = e^{-\sum_{j=1}^{n} \hat{\lambda}_j t_j} \tag{2-16}$$

where

λ^* = allowable system failure rate

$\hat{\lambda}_j$ = allocated failure rate for the jth unit in the system

One approach to allocating system reliability requirement to the constituent units is to base the allocation on the available failure data of existing equipment which is comparable in function, complexity, and application. From the available reliability data on the unit of interest, the unit failure index for each constituent unit in the system is obtained in consideration of unit failure rates, duty cycle, application stress level, and environmental conditions. Allocation of allowable system failure rate is then made by using an allocation index the ratio of unit failure index to total system failure index. Thus, if a given unit now constitutes A per cent of the total system failure index, it is reasoned that the same unit should not be permitted to exceed A per cent of the total system failure index under the new allocation.

Following the line of the above arguments, allocated failure rate for the jth unit is obtained from the relation

$$\frac{\hat{\lambda}_j \, t_j}{\lambda^* t} = \frac{W_j}{\sum_{j=1}^{n} W_j} = A_j \tag{2-17}$$

where

W_j = unit failure index of the jth unit

A_j = allocation index of the jth unit, and $\sum_j A_j = 1.0$,

which now leads to the relation

$$R^*(t) = e^{-\lambda^* t} = e^{-\hat{\lambda}_1 t_1} \cdot e^{-\hat{\lambda}_2 t_2} \cdots e^{-\hat{\lambda}_n t_n} \tag{2-18}$$

$$= e^{-A_1 \lambda^* t} \cdot e^{-A_2 \lambda^* t} \cdots e^{-A_n \lambda^* t}$$

Since $R^*(t) = \prod_{j=1}^{n} \hat{R}_j(t_j)$, the allocated reliability for the jth unit for t_j operating period is determined from

$$\hat{R}_j(t_j) = e^{-A_j \lambda^* t} = [R^*(t)]^{A_j} \tag{2-19}$$

Then, the allocated failure rate and mean time to failure of the jth unit are respectively

$$\hat{\lambda}_j = -\frac{\log \hat{R}_j(t_j)}{t_j} \tag{2-20}$$

$$\hat{\theta}_j = -\frac{t_j}{\log \hat{R}_j(t_j)} \tag{2-21}$$

Or, using the relation in (2-19) that $\log \hat{R}_j(t_j) = A_j \log R^*(t)$ gives

$$\hat{\lambda}_j = -\frac{A_j \log R^*(t)}{t_j} \tag{2-22}$$

$$\hat{\theta}_j = -\frac{t_j}{A_j \log R^*(t)} \tag{2-23}$$

In determining the system's reliability which gives the probability of the system being able to survive for a given time period of interest, one or more units within the system might sometimes fail without necessarily causing failure of whole system. In other words, one or more unit failures for a given system is not a necessary and sufficient condition for the occurrence of the system failure. If this is the case in a given serial system, the following conditional probability that the system will continue to function satisfactorily despite the occurrence of the particular unit failure(s) is defined to reflect the unit importance in the system function.

Let E_j be the probability that the system continues to function satisfactorily under the condition that jth unit has failed, but that all other units in the system are operable. Then, under this modified condition, the probability of system success in the presence of jth unit failure over a time period of t_j hours will be given by the sum of the probabilities of following two mutually exclusive events:

(1) no failure of the jth unit occurs: $R_j(t_j)$

(2) failure of the jth unit occurs but there is no system failure: $[1 - R_j(t_j)] \cdot E_j$

Hence, this probability is

$$R_j(t_j) + [1 - R_j(t_j)]E_j \tag{2-24}$$

Then, the reliability of the system over t system operating hours is given by

$$R(t) = \prod_{j=1}^{n} \{R_j(t_j) + [1 - R_j(t_j)] \cdot E_j\} \tag{2-25}$$

assuming that the unit independence of failure again holds as before, and that all E_j factors for multiple unit failures are multiplicative (i.e., the probability of the system success given that the jth unit and Kth unit failures is assumed to be $E_j \cdot E_K$).

Given a system reliability requirement $R^*(t)$, the allocations among the units are accomplished through the use of allocation index as before. Thus $R_j(t_j)$ can be determined so that the relation established in (2-19) again holds in the allocation process.

Hence

$$\hat{R}_j(t_j) + [1 - \hat{R}_j(t_j)] \cdot E_j = [R^*(t)]^{A_j} \tag{2-26}$$

After rearranging the terms, $\hat{R}_j(t_j)$ is given by

$$\hat{R}_j(t_j) = \frac{[R^*(t)]^{A_j} - E_j}{1 - E_j} \qquad (2\text{-}27)$$

With the approximation $\hat{R}_j(t_j) = e^{-\hat{\lambda}_j t_j} \approx 1 - \hat{\lambda}_j t_j$, where $\hat{\lambda}_j t_j$ is small, the approximated allocated failure rate and mean time to failure are respectively

$$\hat{\lambda}_j = - \frac{A_j \log R^*(t)}{(1 - E_j)\, t_j} \qquad (2\text{-}28)$$

$$\hat{\theta}_j = - \frac{(1 - E_j)\, t_j}{A_j \log R^*(t)} \qquad (2\text{-}29)$$

Here approximation is carried out twice in opposite direction so that errors cancel in part. Equation (2-27) is the equivalent allocation Equation (1-3) derived in the AGREE report.[17] Determination of the E_j factors in a given situation requires both a good knowledge of the unit capability, and of the operational requirements of the system. Certainly, any reasonable estimation of these factors during the early formative stage of a system design is much dependent upon the subjective analysis of the operational requirements for the particular unit within a system.[18] However, a fairly reliable information required for the quantification of the E_j factors for a given sytem may be obtained from

[17] AGREE Report, *loc. cit.*

[18] For a subjective analysis method of quantifying judgemental value system, see C. W. Churchman, R. L. Ackoff, and E. L. Arnoff, *Introduction to Operations Research* (N.Y.: John Wiley & Sons,Inc.,1957), pp. 136-144.

the existing data of failure reporting procedures currently in use at the field which gives the frequency of unit failures, and its after-effect on the system's performance. In quantifying the E_j factors in a given system, sufficient attention should be given to such factors as the particular structure (physical make-up) of the unit, safety of the operating personnel, and location of the unit which fails since these factors influence the successful system operation should the particular unit fail.

Determination of Allocation Index

A crucial part of the functional allocation method is to obtain the applicable failure index for the allocation units within a system. Such failure indices for the constituent units are necessary to determine the allocation indices which serve to establish the meaningful reliability allocations. The allocation index for the unit is defined in Equation (2-17).

How to obtain reliable data required to compute the allocation index for the unit is a serious question, and sometimes a difficult problem. The unit failure indices (w_j) represent the measures of relative reliabilities of the units to which system reliability requirements being apportioned. The unit failure indices are determined in consideration of base failure rates of the units, application stresses, and environmental factors, and duty cycle of the units.

At the time of preliminary system design, the system designer generally has a good knowledge of the functional equipment which will be required to perform a given system function, and thus, he is able to draw the detailed functional specifications on the units. These preliminary ideas on the equipment and their specifications enable a comparison to be made of the similar equipment for which reliability and usage information has been accumulated.

Many companies have their own data of average failure rates of the various units used in their design work, based on past experience with the particular unit which has been statistically evaluated. Such information is often quite useful for determining unit failure index of the specific unit. The failure-reporting procedures currently in use at the field are also useful for the estimation of the particular equipment failure rate under actual "use" conditions. The accuracy of the estimate is, of course, dependent upon the accuracy of data reporting in the field and the availability of the adequate supporting information.

In some cases, the entire system may stem out of a new concept on which there are no past reliability data for any of its constituent unit. One approach in this case would be to base the estimate of a required unit failure rate on the nearest similar equipment of comparable function and complexity. This approach would be permissible until such time when the design studies disclose differences in such estimates, and then, appropriately "update" the allocations. A still another approach for the estimate of a unit failure rate is to estimate the kind and number of active parts in each functional unit in the system, and associate a known failure rate to each part, and thereby obtain an estimate of the failure rate of the unit.

Generally, a great deal of usage experiences and failure information exist, and have been well documented to be of extensive use for electronic and electromechanical equipment. However, sufficient reliability information has not yet generated and tabulated for the non-electronic type equipment. The reason for this lack of information on the non-electronic equipment lies in the fact that most non-electronic types of equipment are usually designed specifically for the particular system requirements and configurations. This is true, for example, of turbogenerator, pump, valves, and combustion chamber, etc. Thus, the failure information on this type of non-electronic equipment is usually limited and specific to the particular unit, and it does not readily apply to other systems of the same type if there is a difference between two operational conditions. On the other hand, electronic units are frequently standardized items, and thus, they are often interchangeable between two equivalent use conditions.

Generally, electronic or non-electronic units or components are designed to withstand certain nominal operational stresses in operation. Thus, the failure rate figures for a certain unit or component apply to definite operating stress conditions--for example, to an operation at rated voltage, current, frequency, temperature, and at a predicted level of mechanical stresses such as vibration, shock, acceleration, etc.

It is a well known fact that when the operational stresses are increased above the rated level, the unit failure rate increases rather rapidly above the nominal failure rate. Conversely, the unit failure rate decreases when the operational stresses are decreased below the rated level.[19]

Therefore, before a specific use can be made of any available failure rate for a given unit, it is always necessary to adjust it for identifiable differences between two application stresses. These adjustments are in the form of factors called application stress factor, K_S, which multiply the available failure rate of a given unit to "correct" it according to its expected application (electrical, mechanical, and environmental) stresses.

The application stress factor, K_S, is defined to be

$$K_S = \frac{\left[\begin{array}{c}\text{Expected Failure Rate of the Unit Under} \\ \text{its Anticipated Application Stresses}\end{array}\right]}{\left[\begin{array}{c}\text{Base Failure Rate of the Unit Under} \\ \text{Typical Application Stresses}\end{array}\right]} \qquad (2\text{-}30)$$

The K_S factor may be larger or smaller than unity and equal to unity when the base or standard stress conditions for which the typical failure rate is available are approximately equivalent to anticipated application stress conditions.

The prime component of application stresses which act on the unit within a system may be grouped into two categories:

1. Operating stresses which are predominantly present when

---

[19]See, for example, *MIL-HDBK-217, Reliability Stress and Failure Rate Data for Electronic Equipment* (U. S. Government Printing Office, 1962) provides the failure rate versus usage trade-off relationships for various components of the electronic equipment.

the unit is in operation or energized such as voltage, frequency, current, temperature, pressure, or internally-generated heat, and mechanical stresses such as vibration, shock, friction or impact if the unit or any part of the unit has to perform mechanical motions.

2. Internal environmental stresses which are present in a system whether or not the particular unit is actively operating, such as radiation, humidity, atmospheric pressure, corrosion, and the system-generated heat which affects the units in the system.

In addition to the application stresses, the available failure rate of a given unit must be corrected for identifiable differences in gross use (external) environments. By this is meant that particular equipment in question will be operating in certain use environments such as shipborne, fixed ground equipment, manned aircraft, missile equipment, satellite equipment (launch and boost phase) or satellite equipment (orbit or midspace phase), etc.

The factor which accounts for the differences between base or "standard" environment and anticipated use environment is called use environment factor. The general formula used in estimating the use environment factor, K_u, is

$$K_u = \frac{\begin{bmatrix}\text{Expected Failure Rate of the Unit} \\ \text{Under Anticipated Use Environment}\end{bmatrix}}{\begin{bmatrix}\text{Base Failure Rate Available for} \\ \text{Equivalent Unit under Typical Use} \\ \text{Environment}\end{bmatrix}} \tag{2-31}$$

Military Standard 756, "Reliability of Weapons Systems, Procedures for Prediction and Reporting Prediction of," gives the K_u value of 6.5 for airborne environment and 80.0 for missiles environment, where stationary ground equipment is used as the equipment of standard environment.

The third element required to obtain the allocation index for the unit is the duty cycle. The concept of duty cycle must be included in determining the allocation index in order to reflect any variance in unit operational time requirement as compared to the system mission time. The units which operate for a shorter length of time during a given system mission time should have a relatively high allocation than the units which operate all the time. The duty cycle of the jth unit, d_j, is defined as

$$d_j = \frac{t_j}{t} \tag{2-32}$$

where

t_j = required operating hours of the jth unit during t system operating hours

5 = system mission duration requirement

In summary, to obtain the failure index for a particular unit in a system, the following steps are required:

1. Describe the functional configurations and define the functional units and its system.

2. Determine anticipated application stress level and gross use environment under which the unit is expected to achieve the allocated

reliability requirement.

3. Determine unit duty cycle (d_j).

4. Obtain its "Base" failure rate (B_j).

5. Obtain its K_S factor from the appropriate references according to its anticipated stress level.

6. Obtain its K_u factor from the appropriate references.

7. Multiply the number in 3, 4, 5, and 6 to obtain unit failure index; that is

$$W_j = B_j \cdot d_j \cdot K_S \cdot K_u$$

The above process is repeated for all functional units in the system. If W_j is the unit failure index of the jth unit, total system failure index is obtained by the summation of all unit failure indices, and the allocation index for the jth unit is defined in (2-17) as $A_j = W_j / \sum_{j=1}^{n} W_j$, which is the ratio of the unit failure index to the total system failure index.

For a unit which experiences abrupt changes in the application stresses during a given operating time, and consequently assumes the different failure rates during this operating period, an approximate average failure rate to use is already given in Equation (2-14). The situation such as this where the unit failure rate changes because of the changes in the application stresses occurs nearly in all airborne systems in which the failure rates of some units or components may be many times higher during the launching than while cruising, because stress levels change abruptly.

Many individual manufacturers have started to compile the appropriate failure data for their products through their own life testing or keeping logs on their delivered equipment and this information is readily available to the user upon request. Other organizations such as ARINC, and VITRO have been performing extensive reliability studies of equipment in the field for the Armed Services.

Also there exist a great amount of usage experiences and failure information tabulated and graphed for various application stresses and use environments from which system designer can obtain the necessary failure rate at desired operating conditions. Some of these reference sources include the following:

1. MIL-HDBK-217, "Reliability Stress and Failure Rate Data for Electronic Equipment," August 8, 1962, U. S. Government Printing Office, Washington, D. C.

2. RCA Service Company, "The Prediction and Measurement of Air Force Ground Electronic Equipment Reliability," Final Engineering Report, Contract No. AF 30(602)-1623, RADCTN 58-307, August, 1958.

3. ARINC Research Corporation," Improved Techniques for Design Stage Prediction," Vol. I of Air Force Reliability Assurance Program Progress Report No. 2, Contract AF 33 (600)-38438, April, 1959.

4. J. H. Hershey, "Reliability and Maintainability of Military Electronic Equipment," Third Annual Signal Maintenance Symposium, Fort Monmouth, New Jersey, April, 1959.

5. D. E. Johnston, and D. T. Mcruer, "A Summary of Component Failure Rate and Weighting Function Data, and their Use in Systems

Preliminary Design," Wright Air Development Center, WADC TR-57-688, December, 1957.

6. R. L. Vander Hamm, "Component Part Failure Rate Analysis for Prediction of Equipment Mean Life," IRE Convention Record, Part 6, 1958, pp. 72-76.

7. T. H. Bollman, "Instructions and Data for Failure and Prediction," Bell Telephone Laboratories Report, October 11, 1957.

8. RCA Technical Report 59-46-1, "Reliability Stress Analysis for Electronic Equipment," January 15, 1959.

9. Johnston, D. E., K. A. Ferrick, and E. L. Alexander, "A Compilation of Component Field Reliability Data Useful in Preliminary Design," WADD Technical Report 60-330, Supplement 1, Aeronautical Systems Division, Air Force System Command, U. S. Air Force (November, 1961).

In addition to the above references, MIL-STD-210A, "Climatic Extremes for Military Equipment," provides comprehensive, worldwide environmental coverages.

Concluding Remarks

In the preceding sections, it has been pointed out that the reliability allocation of a large-scale system involves numerous complicated problems due to many interacting factors involved in the allocation process. The complexity of the problem varies to the large extent with the specified mission requirements, available knowledge of its feasibility and inherent limitations, and the specific usage experiences obtained from the similar systems previously developed. Because of the complexity of relating many interacting factors in the allocation process, the system reliability allocation is carried out through two distinct methods in this study.

In this chapter, a reliability allocation method has been developed for assigning the specified system reliability requirement into numerical reliability requirements at the major functional equipment groups which comprise a given system. The allocation method seems to have general applications due to the fact that the required input data to the allocation model are based only on the available information during the early design stages of a system development program.

The method developed in this chapter differs from the existing ones in that (a) the method takes into account not only the state-of-the-art, complexity, and duty cycle but also applicational stresses and use environment factors which influence the unit reliability; (b) the reasoning employed in the allocation process is rather general so that the method can be applicable to various kinds of systems (i.e., not necessarily restricted to the electronic system, and at the same time, the method does not require the assumption of equal failure rates among various modules); (c) the method tends to provide more realistic results as the usage experiences accumulate in the field; and (d) the method does include the concept of system effectiveness, because of the E_j factors in Equation (2-25), under the conditions that one or more failures of particular units will not necessarily cause the failure of an entire system.

CHAPTER III

DETAILED ALLOCATION OF SYSTEM RELIABILITY

Preliminary Considerations

The previous chapter has described the functional allocation method through which system reliability requirement or "allowable" system failure rate is apportioned among the functional units within a proposed new system. However, this does not complete the development of a comprehensive allocation model. The reliability allocation through this method is only performed at several major functional equipment levels. Furthermore, the functional allocation process assumes an equality of improvement feasibility for all functional units which may not be true in practice. Thus the allocations arrived at by this method may be considered only as a preliminary one--for use only as an initial basis for specification of reliability requirements at the functional equipment levels.

The application of this method will only yield early decisions at these levels, and must therefore be reviewed and modified early in the design stage as soon as the detailed feasibility study discloses the discrepancies between allocated requirement and improvement feasibility. Sometimes it may turn out, for example, that a five-to-one reduction in the failure rate of a certain unit is entirely feasible, whereas a two-to-one reduction in another unit would be beyond state-of-art capability for some time to come.

If the allocated reliability requirements at the subsystem levels greatly exceed the predicted values obtained through the feasibility study, it is quite logical that a management decision should be made regarding abandonment of the present design and initiating a new one or possibly concentrating the effort on improving the design. In any case, the best decision is primarily a function of improvement feasibility in reliability that can be achieved for a given expenditure of effort, usually within the constraints of available funds and a time schedule.

If the allocated reliability requirements dictate an average improvement in unit reliabilities, areas should be indicated for which more effort is to be concentrated. This is the point where design decisions, for the purpose of increasing the inherent unit reliability and achieving the allocated requirement, must be made as to (a) redundancy versus further derating of the components or units, or (b) redesigning weak units versus a search for high reliability components, or (c) a combination of these approaches.

In designing for the required reliability, the derating of components or units is a standard practice used in the system design--this means operating components or units at only one-half or even less of their rated (nominal) values of temperature, pressure, voltage, wattage, etc. Significant increases in component and unit reliability can be achieved in this way. Sometimes only one of these stress parameters needs to be reduced to bring the component failure rate down to the desired level.

Another powerful technique in achieving the required reliability is the introduction of several redundant subsystems in the form of parallel or stand-by units so that one or more spare units will always be available, and the failed unit can be repaired under less pressing circumstances. Generally the use of redundancy to achieve the required reliability goal is the quickest and easiest approach if the time involved is of prime importance, and also the cheapest approach if the addition of several components is more economical in comparison with the cost of redesign which inevitably requires more basic research on the target units.

However, in weighing its disadvantage, redundancy will prove too expensive if the components are costly, and inevitably exceed the limitations on weight, size, power requirements, and possibly increase maintenance effort as well. Redundancy also requires sensing and switching devices which are by no means perfectly reliable, and thus may offset the advantage of employing redundancy.

Because the reliability of a serial system is determined by the sum of the component failure rates, design simplifications which involve a substantial reduction of the number of components used in the equipment is another important approach toward reliability design. To design one highly reliable unit with the minimum possible number of components is a sound and formal design approach for achieving the reliability requirement.

Thus, in view of the above several promising techniques for achieving the allocated reliability requirements for the units, the system designer must make trade-off analysis to select an optimum design policy for attaining a needed level of reliability at the least expenditure of money, personnel and development time. An optimum decision can always be made most effectively if there exists the information which shows the changes in system reliability resulting from various combinations of inherent unit reliability and redundancy, and their associated costs.

To summarize the above, the allocation of system reliability requirement during the initial phases of system design through the use of functional allocation method permits us to establish a first set of reliability requirements at the major sybsystem level, which should be reviewed against the predicted feasibility within the constraints of the state-of-the-art, development costs, and time.

There are several promising techniques to cope with achieving the required level of reliability when the allocated requirement for a certain unit exceeds the predicted value. Any choice of these design alternatives to achieve the needed level of reliability must be based on the sound trade-off analysis which serves to select one optimum policy consistent with the concept of system effectiveness.

The detailed allocation method of system reliability is designed to select the optimum solution in the context of this trade-off analysis. This chapter gives the detailed development of this allocation method, starting with the mathematical formulation of the allocation problem and the detailed computational procedures for the solution, and finally including some numerical results.

Mathematical Formulation of the Reliability Allocation Problem

The foregoing discussions on the necessity of developing the detailed reliability allocation model lead to the following mathematical formulation of the problem. To formulate the problem mathematically, it is assumed that the system under consideration consists of n functional units which are connected in series, and failure of any unit in the system is assumed to be an independent event. Associated with the n functional units there exist several choices of design alternatives that the system designer can employ in order to meet the allocated reliability requirement.

Let a_K represent the design alternatives available for the Kth unit with a specified inherent unit reliability, and let $R_K(m_K, a_K)$ denote the known reliability function of the Kth unit when m_K identical units of design alternative a_K are used in redundancy. It is assumed that $C_K(m_K, a_K)$ and $W_K(m_K, a_K)$ denote respectively the cost and weight function for the design alternative a_K using m_K identical units in the Kth functional unit in the system.

Given the overall restrictions on the system cost of C and weight of W, the problem is to determine which design alternative to select with the specified level of inherent unit reliability, and how many redundant units to use which will result in the greatest reliability while keeping the total system cost and weight within the allowable amounts. This formulation of the problem leads to the maximization of system reliability R given by the product of unit reliabilities

$$R = \prod_{K=1}^{n} R_K(m_K, a_K) \tag{3-1}$$

subject to the constraints

$$\left.\begin{aligned} &\sum_{K=1}^{n} C_K(m_K, a_K) \le C \\ &\sum_{K=1}^{n} W_K(m_K, a_K) \le W \end{aligned}\right\} \tag{3-2}$$

where

$$m_K = 1,2,3,\ldots; \qquad a_K = 1,2,3,\ldots,A_K;$$

and

$$C_K(m_K, a_K), \quad W_K(m_K, a_K) > 0 \quad \text{for}$$

$$K = 1,2,3,\ldots,n.$$

It is to be noted that the mathematical problem formulated above can be thought of as an n-stage decision problem, where at each stage K the decisions are made on how many redundant units to use and what type of design alternative to select, that is, m_K and a_K are selected. The basic nature of this allocation problem remains unchanged regardless of how many functional units there exist in the system; in other words, it has the same functional form regardless of what n happens to be for a given problem. Note also that this problem involves two constraints over which the decision variables are to be optimized, and also requires the decision variables to be integral values.

There are two conventional methods of solving this type of optimization problem; direct enumeration methods, and calculus method. Direct enumeration method is that of simply enumerating all possible sets of nonnegative integers m_K and a_K which satisfy the given constraints, evaluating the objective function for each set, and then selecting that set or sets of m_K and a_K which yield the maximum value of the objective function. Even for a simple situation involving only ten units in the system with two design alternatives for each unit, and each independent variable m_K can take on five different values for each alternative for which cost and weight constraints are satisfied, there exist at least $(2 \times 5)^{10}$ different candidate sets to be investigated before an optimal solution is found. It would take more than 2700 hours to investigate them at the rate of one set of m_K per millisecond. Thus, this approach is not only unrealistic but practically impossible as a solution to the problem.

On the other hand, there are many difficulties encountered in treating this problem strictly from the point of view of calculus. The method of calculus is directed toward the optimization problem which involves continuous variation of the independent variables. When the optimal m_K are large, it is sometimes satisfactory to treat m_K as continuous variables and then round off the solutions so obtained to the nearest integers satisfying the constraints. However, the accuracy of the solution so obtained is seriously affected by this round-off when the m_K are small in a given situation. Even if $R_K(m_K,a_K)$, $C_K(m_K,a_K)$, and $W_K(m_K,a_K)$ were continuously differentiable in m_K, the usual method of calculus does not easily render a solution to the problem because the existence of design alternatives a_K causes discontinuity in the above functions.

In summary, then, two usual techniques for solving the optimization problem posed here will not, in general, provide an optimal solution to the reliability allocation problem.

One powerful approach to this type of optimization problem would be to employ the dynamic programming technique. Dynamic programming approach as will be seen later will readily yield an exact solution to the problem in (3-1), and indeed yield all alternative solutions if the optimal solution is not unique. The basic characteristic of the dynamic programming approach is, then, that aspect of the computational procedure which makes it possible to solve a given n-stage decision problem by solving a whole sequence of problems, starting with one-stage, then a two-stage on up to finally the n-stage problem. What makes this approach feasible in practice is the fact that one stage can be added at a time, and the solution for K-stage problem is obtained easily from the solution for K-1 stages by adding the Kth stage and making use of the solution for K-1 stages provided that the solution for K-1 stages is known for all possible quantities of the cost and weight which may be available for allocation to the first K-1 stages.

As will be seen, the computational efficiency of the dynamic programming approach over the direct enumeration method lies in the fact that at each stage for given quantities of cost and weight, all non-maximizing combinations of decision variables m_K and a_K are eliminated and are not carried over to the next stage where they would otherwise have to be reconsidered. Thus, for the same example discussed above, the dynamic programming approach only requires $(2 \times 5)10$ candidate sets to investigate rather than $(2 \times 5)^{10}$ sets required in the direct enumeration method.

The saving in the computational effort becomes enormous as the number of stages increases for a given problem. Furthermore, a remarkable savings in time results from the fact that the dynamic programming approach requires only a single multiplication in computing the numbers to be compared at each stage, whereas n multiplications are required to evaluate the objective function for a given set of decision variables in the direct enumeration method.

Although it is sometimes helpful to begin solving the problem with the explicit analytical form of the objective function and of the constraint functions, it is by no means necessary to try to approximate the analytic expression from the data when this input information is actually available as a table. The dynamic programming approach in this context is just as suitable when input data are available in the tabular form. In fact, the availability of input data on cost and weight as a straight table is common in most cases, and generally represents the existing conditions much better.

One can thus assert, based on the nature of the reliability allocation problem, that the technique of dynamic programming yields a simple, quick, and yet accurate solution to the problem considered here. However, there are difficulties which arise in using the dynamic programming approach due to the fact that the problem has two constraints over which a set of decision variables must be maximized. It would soon become clear that as the number of constraints in a given problem increases, the computation required for solving the problem increases exponentially.

At this point, it might appear that the difficulties involved in solving the present problem would not be much greater than those encountered in solving the problem having a single constraint. However, this is not at all true. For a fairly large-scale problem, it is sometimes impossible to solve the problem even on the high-speed digital computer. In order to overcome these computational difficulties in the present problem, an attempt is made to reduce the dimensionality of the problem from two to one through the introduction of the Lagrange multiplier, and accordingly the present problem is modified in the next section.

Reduction in Dimensionality

In place of the problem of maximizing the function R in (3-1), constrained by the relations (3-2), one may consider the problem of maximizing the modified function

$$R_1 = \left[\prod_{K=1}^{n} R_K(m_K,a_K)\right] \cdot e^{-\lambda \sum_{K=1}^{n} W_K(m_K,a_K)} \tag{3-3}$$

subject to the constraint

$$\sum_{K=1}^{n} C_K(m_K,a_K) \leq C \tag{3-4}$$

where

$$m_K = 1,2,3,\ldots,; \qquad a_K = 1,2,3,\ldots,A_K;$$

$$C_K(m_K,a_K) > 0 \quad \text{for} \quad K = 1,2,3,\ldots,n$$

and λ is a nonnegative parameter called Lagrange multiplier.[1]

The above modified problem is more easily solved than the original problem posed in (3-1) and (3-2), since there is only a single constraint that a set of the maximizing values of m_K and a_K must satisfy. The modified formulation of the problem permits the reduction in dimensionality of the problem by one because the weight constraint in (3-2) no longer appears as a constraint in the modified

---

[1]The use of Lagrange Multiplier is suggested in the paper by Richard E. Bellman, "Dynamic Programming and Lagrange Multiplier," *Proceedings of the National Academy of Science, U.S.A.*, Vol. 42 (1956), pp. 767-769.

problem, but only appears in the objective function.

The problem with a large dimension often exceeds computer memory capacity and demands excessive time to carry out the necessary computations in the course of obtaining an optimal solution. The modified formulation by introducing the appropriate number of Lagrange multipliers actually reduces the computational problem of large size to a manageable size in many cases.

The use of the Lagrange multiplier in the above formulation actually leads to assign a penalty in the form of system unreliability to the amount of weight used, and removing the weight constraint from the problem. One may first imagine that there is no constraint on the amount of weight used, but that each unit of weight used for redundancy results in λ units of unreliability being added to the overall system reliability. It is then clear that if λ is zero, one can use any quantity of weight in an unlimited manner, while if λ is quite large, one may not use a quantity of weight at all because the resulting penalty is too costly.

Each choice of λ will result in some value for the exponent of the exponential term in (3-3), $\sum_{K=1}^{n} W_K(m_K^*, a_K^*)$ where $m_K^* = m_K^*(\lambda)$, and $a_K^* = a_K^*(\lambda)$ are the maximizing values of the m_K and a_K for a given value of λ. If one can determine λ such that the relation $\sum W_K(m_K^*, a_K^*) = W$ holds, as indicated above, one has solved the original two-constraints problem without explicitly introducing the weight constraint.

The equivalence of two formulations can be established by proving that the optimal solution to the modified problem of maximizing the function R_1 in (3-3), satisfying the relation given in (3-4) is in fact an optimal solution to the original problem of maximizing the function R in (3-1) satisfying the two constraints given in (3-2): First assume that $m_K^*(\lambda)$, $a_K^*(\lambda)$, for K = 1,2,...,n is a set of maximizing values obtained for the modified problem in (3-3) and (3-4), and furthermore the following relation

$$\sum_{K=1}^{n} W_K(m_K^*(\lambda), \quad a_K^*(\lambda)) = W \tag{3-5}$$

holds. Note then that the sequence $\{m_K^*(\lambda),\ a_K^*(\lambda)\}$ is in fact a feasible solution to the original problem in (3-1) satisfying two constraints in (3-2). Now suppose that an optimal solution to the original problem in (3-1) and (3-2) is $\{m_K^o, a_K^o\}$ for K = 1,2,3,...,n, and

$$R(M^o,A^o) = \prod_{K=1}^{n} R_K(m_K^o,a_K^o) > \prod_{K=1}^{n} R_K(m_K^*,a_K^*) = R(M^*,A^*) \tag{3-6}$$

where M^o, A^o, M^*, and A^* denote the n-component vector containing the optimal solution, respectively. However, since $\{m_K^*(\lambda),a_K^*(\lambda)\}$ is an optimal solution to the modified problem in (3-3) and (3-4), it follows that

$$R_1(M^*,A^*) = \left[\prod_{K=1}^{n} R_K(m_K^*,a_K^*)\right] \cdot e^{-\lambda \sum_{K=1}^{n} W_K(m_K^*,a_K^*)} \geq \tag{3-7}$$

$$\left[\prod_{K=1}^{n} R_K(m_K^o, a_K^o)\right] \cdot e^{-\lambda \sum_{K=1}^{n} W_K(m_K^o, a_K^o)} = R_1(M^o, A^o)$$

Then, since

$$\sum_{K=1}^{n} W_K(m_K^*, a_K^*) = \sum_{K=1}^{n} W_K(m_K^o, a_K^o) = W,$$

this means

$$R(M^*, A^*) = \prod_{K=1}^{n} R_K(m_K^*, a_K^*) \geq \prod_{K=1}^{n} R_K(m_K^o, a_K^o) = R(M^o, A^o) \qquad (3\text{-}8)$$

This, however, contradicts the relation in (3-6). Therefore $\{m_K^*(\lambda),\ a_K^*(\lambda)\}$ is an optimal solution to the original problem in (3-1) and (3-2).

The above arguments point out the fact that one can simply choose an arbitrary nonnegative λ, find a maximum of the modified (Lagrangian) function having the cost constraint only, and then one has, as a result, a solution to the original problem with two constraints. However, it also indicates nothing about the manner in which one can obtain the maxima of the modified function, but merely asserts that if one can find them, then one can also have maxima of the original problem.

The above observations then suggest the following computational procedures for solving the reliability allocation problem in (3-3) and

(3-4). First begin by selecting an arbitrary nonnegative value of λ, and solve the modified one-dimensional problem in (3-3) and (3-4). If the optimal solutions $\{m_K^*, a_K^*\}$ for $K = 1,2,3,\ldots,n$, found for (3-3) and (3-4) are such that the relation $\sum_{K=1}^{n} W_K(m_K^*, a_K^*) = W$ holds, one has obtained an optimal solution to (3-1) and (3-2). If $\sum_{K=1}^{n} W_K(m_K^*, a_K^*) \neq W$, select a new value of λ, resolving the problem for the new value of λ, and examining the resulting $\sum_{K=1}^{n} W_K(m_K^*, a_K^*)$.

This process is repeated until the new λ yielding $\sum_{K=1}^{n} W_K(m_K^*, a_K^*) = W$ is finally found. Usually several iterations will suffice to obtain this desired result, depending upon the effort expended in selecting a new λ at each iteration. One may use the two values of λ and two resulting values of $\sum W_K(m_K^*, a_K^*)$ to interpolate or extrapolate linearly to estimate a new trial value of λ which should produce $\sum W_K(m_K^*, a_K^*) = W$. One then resolves the problem using this new value of λ. One may plot the resulting three points and thus select a new value of λ by connecting these points.

It is to be noted that for the case that m_K must be integral numbers such as the case considered here, it is not always true that one can find the value of λ yielding $\sum W_K(m_K^*, a_K^*) = W$, since both constraints do not necessarily hold as a strict equality. However, one may proceed in this case by varying λ to make a value of the objective function as large as possible while not exceeding either of the constraints.

From a practical point of view, to obtain an optimal solution, one determines by trial and error a value of λ such that the eliminated constraint holds as a strict equality. Then, from what has been proved above one knows that this method has indeed provided an optimal solution to the original problem.

It would be noteworthy that in most cases of system reliability allocation problems, the system designer is not only interested in achieving the maximum system reliability for some given level of weight allowance, but rather in exploring the entire range of system reliabilities obtained as a function of the various amount of allowable system weights. If this is the case in a given situation, no computational efforts are wasted during the search of correct value λ. Since each value of λ selected entails a complete solution of the original problem, the above method will indeed produce a spectrum of this desired information in the course of determining the correct value of λ which ultimately leads to an optimal solution satisfying the weight constraint.

As has been pointed out, the use of Lagrange multiplier is intimately related to the system unreliability resulting from the restriction on the system weight. This interpretation of the Lagrange multiplier would point out an interesting fact that if an initial guess at λ, the resulting value of $\sum W_K(m_K^*,a_K^*)$ is greater than W, then the value of λ in the next trial must be increased so as to decrease the total system weight which will in turn reduce the maximum system reliability achievable.

The above logical observation that as the value of the Lagrange multiplier increases, the resulting system weight and system reliability will decrease can be verified as follows:

Again let $M^* = (m_1^*, m_2^*, \ldots, m_n^*)$, and $A^* = (a_1^*, a_2^*, \ldots, a_n^*)$ be a vector containing a set of maximizing values for the objective function in (3-3) subject to the constraint in (3-4), respectively. Assume that λ and μ are two particular values of the Lagrange multiplier that produce the solution vectors $\{m_K^*(\lambda), a_K^*(\lambda)\}$, $\{m_K^*(\mu), a_K^*(\mu)\}$ for $K = 1,2,\ldots,n$, and furthermore it is assumed that the relation

$$\sum_{K=1}^{n} W_K(m_K^*(\lambda), a_K^*(\lambda)) > \sum_{K=1}^{n} W_K(m_K^*(\mu), a_K^*(\mu))$$

holds.

Since by assumption $\{m_K^*(\lambda), a_K^*(\lambda)\}$ is the solution vector produced by λ, $\{m_K^*(\lambda), a_K^*(\lambda)\}$ maximizes the modified (Lagrangian) function R_1 for a given λ, which then implies that the relation

$$R_1(M^*(\lambda), A^*(\lambda)) = R(M^*(\lambda), A^*(\lambda)) \cdot e^{-\lambda \sum_{K=1}^{n} W_K(m_K^*(\lambda), a_K^*(\lambda))} \geq$$

$$R(M^*(\gamma), A^*(\gamma)) \cdot e^{-\lambda \sum_{K=1}^{n} W_K(m_K^*(\gamma), a_K^*(\gamma))}$$

holds for all choices of $M^*(\gamma)$ and $A^*(\gamma)$ including $M^*(\mu)$ and $A^*(\mu)$. Thus, it follows that

$$R(M^*(\lambda), A^*(\lambda)) \cdot e^{-\lambda \sum_{K=1}^{n} W_K(m_K^*(\lambda), a_K^*(\lambda))} \geq$$

$$R(M^*(\mu),A^*(\mu)) \cdot e^{-\lambda \sum_{K=1}^{n} W_K(m_K^*(\mu),a_K^*(\mu))}$$

or

$$\frac{R(M^*(\lambda),A^*(\lambda))}{R(M^*(\mu),A^*(\mu))} \geq e^{\lambda\left\{\sum_{K=1}^{n} W_K(m_K^*(\lambda),a_K^*(\lambda)) - \sum_{K=1}^{n} W_K(m_K^*(\mu),a_K^*(\mu))\right\}} \tag{3-9}$$

In the same way, upon interchanging the role of $M^*(\lambda)$, $A^*(\lambda)$ by $M^*(\mu)$ and $A^*(\mu)$, one obtains the relation

$$\frac{R(M^*(\lambda),A^*(\lambda))}{R(M^*(\mu),A^*(\mu))} \leq e^{\mu\left\{\sum_{K=1}^{n} W_K(m_K^*(\lambda),a_K^*(\lambda)) - \sum_{K=1}^{n} W_K(m_K^*(\mu),a_K^*(\mu))\right\}} \tag{3-10}$$

Taking the logarithm of the both sides of inequality relation in (3-9) and (3-10), and noting the assumption that $\sum W_K(m_K^*(\lambda),a_K^*(\lambda)) > \sum W_K(m_K^*(\mu),a_K^*(\mu))$, then yields

$$\lambda \leq \frac{\log R(M^*(\lambda),A^*(\lambda)) - \log R(M^*(\mu),A^*(\mu))}{\sum_{K=1}^{n} W_K(m_K^*(\lambda),a_K^*(\lambda)) - \sum_{K=1}^{n} W_K(m_K^*(\mu),a_K^*(\mu))} \leq \mu \tag{3-11}$$

since λ and μ are nonnegative real numbers, the relation

$$R(M^*(\lambda),A^*(\lambda)) \geq R(M^*(\mu),A^*(\mu))$$

holds if

$$\sum_{K=1}^{n} W_K(m_K^*(\lambda), a_K^*(\lambda)) > \sum_{K=1}^{n} W_K(m_K^*(\mu), a_K^*(\mu))$$

or conversely if $o < \lambda < \mu$, then

$$\sum_{K=1}^{n} W_K(m_K^*(\lambda), a_K^*(\lambda)) > \sum_{K=1}^{n} W_K(m_K^*(\mu), a_K^*(\mu))$$

The above argument shows that given two optimal solutions corresponding to two different values of Lagrange multiplier, the ratio of logarithmic change in the system reliability to the change in the weight expenditure is bounded between the two associated Lagrange multipliers λ and μ. This monotonicity in the value of Lagrange multiplier clearly indicates the direction to move in the next trial in order to find the correct value of λ which ultimately yields the optimal solution to the problem.

Dynamic Programming Approach

In the preceding section, it has been asserted that the reliability allocation problem under consideration can best be handled by the dynamic programming technique. The mathematical problem that arises in the choice of the most efficient redundancy and design alternative achieving the maximum system reliability is that of maximizing the function of 2n variables given in Equation (3-1) subject to the relation in (3-2).

This allocation problem is called a two-dimensional problem since the decision variables of interest at each stage must be maximized over the two-dimensional region defined by (3-2). Since the restriction on the memory capacity of high-speed digital computer is such that it would be more preferable to carry out a large number of one-dimensional problems than one multi-dimensional problem, instead of maximizing the function in (3-1) subject to the constraints in (3-2), the modified formulation of the problem gives the following maximization problem:

Maximize the function defined by

$$R_n = \prod_{K=1}^{n} \left[R_K(m_K, a_K) \cdot e^{-\lambda \cdot W_K(m_K, a_K)} \right] \tag{3-12}$$

subject to the constraint

$$\sum_{K=1}^{n} C_K(m_K, a_K) \leq C \tag{3-13}$$

where $m_K = 1,2,3,\ldots;$ $\qquad a_K = 1,2,\ldots,A_K;$

$$C_K(m_K, a_K), \quad W_K(m_K, a_K) > 0$$

$$K = 1,2,3,\ldots,n,$$

and λ is the Lagrange multiplier, to be determined such that

$$\sum_{K=1}^{n} W_K(m_K,a_K) = W \tag{3-14}$$

for the maximizing values of m_K and a_K.

In developing a computational procedure for solving the above problem it is absolutely essential that the procedure must be capable to examine every set of feasible m_K and a_K while not violating the constraint defined by (3-13).

Since the maximum of $R(m_1,m_2,\ldots,m_n;\ a_1,a_2,\ldots,a_n)$ over the defined region will depend upon n and C, denote by $f_n(C)$ the absolute maximum of R_n in (3-12). Then

$$f_n(C) = \max_{\substack{m_1,\ldots,m_n \\ a_1,\ldots,a_n}} \left[\prod_{K=1}^{n} R_K(m_K,a_K) \cdot e^{-\lambda \cdot W_K(m_K,a_K)} \right] \tag{3-15}$$

where the maximization is taken over the nonnegative integers m_K and a_K satisfying the relation

$$\sum_{K=1}^{n} C_K(m_K,a_K) \leq C \tag{3-16}$$

Then the function $f_n(C)$ gives the maximum system reliability achieved from n serial units when the total amount C of the system cost is available for allocation to these units.

Now suppose that the maximization process in (3-15) proceeds as follows: First select a value of m_n and a_n, respectively, and holding

m_n and a_n fixed, maximize the function R_n over the remaining variables $m_1, m_2, \ldots, m_{n-1}$, and $a_1, a_2, \ldots, a_{n-1}$. The values of m_K, a_K for $K = 1,2,\ldots,n-1$, which maximize the function R_n under these conditions will, of course, depend upon the values of m_n and a_n already selected. Suppose that this maximization is performed for every allowable value of m_n and a_n. Then, obviously, the maximum of the function $f_n(C)$ will be the largest of all the R_n values so obtained, and one thus finds m_K, a_K, $K = 1,2,\ldots,n$ which maximize the function.

Following the above steps, first select a pair of values for m_n and a_n and compute the quantity

$$R_n(m_n,a_n)\, e^{-\lambda \cdot W_n(m_n,a_n)} \max_{\substack{m_1,\ldots,m_{n-1} \\ a_1,\ldots,a_{n-1}}} \left[\prod_{K=1}^{n-1} R_K(m_K,a_K) \cdot e^{-\lambda \cdot W_K(m_K,a_K)} \right] \tag{3-17}$$

The term $R_n(m_n,a_n)e^{-\lambda W_n(m_n,a_n)}$ can be factored out, since this term is no longer dependent upon $\{m_K,a_K\}$, $K = 1,2,\ldots,n-1$. Once m_n and a_n have been chosen, the remaining set of variables $\{m_K,a_K\}$ for $K = 1,2,\ldots,n-1$ must then be maximized over the region defined by

$$\sum_{K=1}^{n-1} C_K(m_K,a_K) \le C - C_n(m_n,a_n) \tag{3-18}$$

where $\quad m_K = 1,2,\ldots; \qquad a_K = 1,2,\ldots,A_K;$

$\quad K = 1,2,\ldots,n-1.$

But the quantity

$$\max_{\substack{m_1,\ldots,m_{n-1}\\ a_1,\ldots,a_{n-1}}} \left[\prod_{K=1}^{n-1} R_K(m_K,a_K) \cdot e^{-\lambda \cdot W_K(m_K,a_K)} \right]$$

for nonnegative integers satisfying the above constraint in (3-18) depends on the values of m_n and a_n chosen, or more specifically, on the quantity $C - C_n(m_n,a_n)$. Thus, one can write

$$f_{n-1}(C - C_n(m_n,a_n)) = \max_{\substack{m_1,\ldots,m_{n-1}\\ a_1,\ldots,a_{n-1}}} \left[\prod_{K=1}^{n-1} R_K(m_K,a_K) \cdot e^{-\lambda \cdot W_K(m_K,a_K)} \right] \quad (3\text{-}19)$$

where the maximization is over the nonnegative integers $\{m_K,a_K\}$, $K = 1,2,\ldots,n-1$, satisfying (3-18).

If one can compute $f_{n-1}(C - C_n(m_n,a_n))$ for every allowable value of m_n and a_n, then it is clear that an optimal choice of m_n and a_n is the one which maximizes (3-17). Then it becomes now obvious that the following recurrence relation

$$f_n(C) = \max_{\substack{C_n(m_n,a_n) \le C\\ a_n = 1,2,\ldots,A_n\\ m_n = 1,2,\ldots}} \left[R_n(m_n,a_n) \cdot e^{-\lambda W_n(m_n,a_n)} \cdot f_{n-1}(C-C_n(m_n,a_n)) \right] \quad (3\text{-}20)$$

holds for $n = 2,3,4,\ldots$.

In the particular case of interest, it is clear that

$$f_n(0) = 0 \quad \text{for} \quad n = 1,2,\ldots . \tag{3-21}$$

and

$$f_1(C) = \max_{\substack{C_1(m_1,a_1) \le C \\ a_1 = 1,2,\ldots,A_1 \\ m_1 = 1,2,\ldots}} \left[R_1(m_1,a_1) \cdot e^{-\lambda \cdot W_1(m_1,a_1)} \right] \tag{3-22}$$

holds for $C \ge C_1(m_1,a_1)$.

The above recurrence relation connecting $f_n(C)$ and $f_{n-1}(C)$ for arbitrary n and C is an immediate consequence of Bellman's principle of optimality which states that an optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.[2]

The recurrence relation of (3-20) provides a theoretical method for obtaining the sequence $\{f_K(C)\}$, $K = 2,3,4,\ldots,n$, inductively, once $f_1(C)$ is computed. To evaluate the maximum in (3-20), one must first evaluate the quantity

$$R_n(m_n,a_n) \cdot e^{-\lambda \cdot W_n(m_n,a_n)} \cdot f_{n-1}(C - C_n(m_n,a_n)) \tag{3-23}$$

---

[2]Richard E. Bellman, *Dynamic Programming* (Princeton, N.J.: Princeton University Press, 1957), p. 83.

for every possible combination of m_n and a_n, and select the largest so obtained. Then, one can simultaneously determine m_n^* and a_n^* which are the optimal values of m_n and a_n, respectively. Thus, if the function $f_{n-1}(C - C_n(m_n,a_n))$ were known, one could reduce the problem of obtaining $f_n(C)$ to a maximization over just a single set of variables m_n and a_n.

Now, to compute $f_{n-1}(C - C_n(m_n,a_n))$, first note that $f_{n-1}(C - C_n(m_n,a_n))$ is defined by (3-19), where the maximization is taken over the nonnegative region defined by (3-18). For any arbitrary nonnegative integer α, let

$$f_{n-1}(\alpha) = \max_{\substack{m_1,\ldots,m_{n-1} \\ a_1,\ldots,a_{n-1}}} \left\{ \prod_{K=1}^{n-1} R_K(m_K,a_K) \cdot e^{-\lambda \cdot W_K(m_K,a_K)} \right\} \tag{3-24}$$

where the maximization is carried out over nonnegative integers satisfying the relation

$$\sum_{K=1}^{n-1} C_K(m_K,a_K) \leq \alpha \tag{3-25}$$

But, if one proceeds just as above, it is clear that

$$f_{n-1}(\alpha) = \max_{\substack{C_{n-1}(m_{n-1},a_{n-1}) \leq \alpha \\ a_{n-1} = 1,2,\ldots,A_{n-1} \\ m_{n-1} = 1,2,\ldots}} \left\{ R_{n-1}(m_{n-1},a_{n-1}) \cdot e^{-\lambda \cdot W_{n-1}(m_{n-1},a_{n-1})} \right. . \tag{3-26}$$

$$f_{n-2}(\alpha - C_{n-1}(m_{n-1}, a_{n-1}))\Bigg\}$$

where

$$f_{n-2}(\beta) = \max_{\substack{m_1,\ldots,m_{n-2} \\ a_1,\ldots,a_{n-2}}} \left\{ \prod_{K=1}^{n-2} R_K(m_K, a_K) \cdot e^{-\lambda \cdot W_K(m_K, a_K)} \right\} \tag{3-27}$$

and the maximization is now over nonnegative integers $m_1, m_2, \ldots, m_{n-2}$, $a_1, a_2, \ldots, a_{n-2}$ satisfying the constraint given by

$$\sum_{K=1}^{n-2} C_K(m_K, a_K) \leq \beta \tag{3-28}$$

Therefore, if one knows the function $f_{n-2}(\beta)$, then $f_{n-1}(\alpha)$can be evaluated for every value of α by carrying out the maximization over a set of the variables m_{n-1}, and a_{n-1}.

Obviously, it is nearly impossible to evaluate all values of the function $f_n(C)$ as C assumes all non-negative values. Thus, it is to be agreed at this point that the values of C and $C_K(m_K, a_K)$ only assume integers, and thus the element of the sequence $\{f_K(C)\}$ will be evaluated at each of these points and only these points.

It is to be noted that to evaluate $f_{n-1}(\alpha)$ for each different value of α, one must repeat the maximization computation in (3-26). The computational process used above can be extended to evaluate $f_{n-2}(\beta)$, and is to be continued until finally one computes at the last step the quantity

$$f_1(\gamma) = \max_{\substack{m_1 \\ a_1}} \left\{ R_1(m_1,a_1) \cdot e^{-\lambda \cdot W_1(m_1,a_1)} \right\} \tag{3-29}$$

where the variables of interest are constrained by the relation

$$C_1(m_1,a_1) \leq \gamma$$

and

$$m_1 = 1,2,3,\ldots; \quad a_i = 1,2,3,\ldots,A_1$$

However, in actually solving the problem, the computational procedure would proceed by determining $f_1(\gamma)$ first, and then compute the $f_2(\eta)$, and finally $f_{n-1}(\alpha)$ and $f_n(C)$. The detailed procedural steps for the numerical solution of the reliability allocation problem is now formulated as follows:

First introduce the sequence of functions defined by

$$f_K(\beta) = \max_{\substack{m_1,\ldots,m_K \\ a_1,\ldots,a_K}} \left\{ \prod_{J=1}^{K} R_j(m_j,a_j) \cdot e^{-\lambda \cdot W_j(m_j,a_j)} \right\} \tag{3-30}$$

where the maximization is carried out over the region defined by

$$\sum_{j=1}^{K} C_j(m_j,a_j) \le \beta \tag{3-31}$$

where $m_j = 1,2,\ldots,;$ $a_j = 1,2,\ldots,A_j;$ $J = 1,2,3,\ldots,K$.

Once $f_1(\beta)$ has been obtained directly from (3-29), the remaining $f_K(\beta)$ can be computed by using the recurrence relation

$$f_K(\beta) = \max_{\substack{m_K \\ a_K}} \left\{ R_K(m_K,a_K) \cdot e^{-\lambda \cdot W_K(m_K,a_K)} \cdot f_{K-1}(\beta - C_K(m_K,a_K)) \right\} \tag{3-32}$$

for $K = 2,3,4,\ldots,n$, and finally one can obtain $f_n(C)$.

In order to compute $f_K(\beta)$ in (3-32), it is to be noted that the maximization must be carried out over every possible combination of m_K and a_K when m_K and a_K are restricted to nonnegative integers such that $C_K(m_K,a_K) \le \beta$. But, observing that the decision variable a_K can only take on the values of $1,2,3,\ldots,A_K$, it is clear that $f_K(\beta)$ in (3-32) can be written as:

$$f_K(\beta) = \max_{\substack{C_K(m_K,a_K) \le \beta \\ a_K = 1,2,3,\ldots,A_K \\ m_K = 1,2,3,\ldots}} \left\{ R_K(m_K,a_K) \cdot e^{-\lambda \cdot W_K(m_K,a_K)} \cdot f_{K-1}(\beta - C_K(m_K,a_K)) \right\}$$

$$= \max \begin{bmatrix} \max\limits_{\substack{C_K(m_K,1) \le \beta \\ m_K = 1,2,3,\ldots}} \left\{ R_K(m_K,1) \cdot e^{-\lambda \cdot W_K(m_K,1)} \cdot f_{K-1}(\beta - C_K(m_K,1)) \right\} \\ \max\limits_{\substack{C_K(m_K,2) \le \beta \\ m_K = 1,2,3,\ldots}} \left\{ R_K(m_K,2) \cdot e^{-\lambda \cdot W_K(m_K,2)} \cdot f_{K-1}(\beta - C_K(m_K,2)) \right\} \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \max\limits_{\substack{C_K(m_K,A_K) \le \beta \\ m_K = 1,2,3,\ldots}} \left\{ R_K(m_K,A_K) \cdot e^{-\lambda \cdot W_K(m_K,A_K)} \cdot f_{K-1}(\beta - C_K(m_K,A_K)) \right\} \end{bmatrix} \qquad (3\text{-}33)$$

which reduces to the problem of finding a maximum among A_K maximums obtained for a given value of a_K. Then it is also clear that $f_1(\beta)$ in (3-29) can be written as

$$f_1(\beta) = \max \begin{bmatrix} \max\limits_{\substack{C_1(m_1,1) \le \beta \\ m_1 = 1,2,3,\ldots}} \left\{ R_1(m_1,1) \cdot e^{-\lambda \cdot W_1(m_1,1)} \right\} \\ \max\limits_{\substack{C_1(m_1,2) \le \beta \\ m_1 = 1,2,3,\ldots}} \left\{ R_1(m_1,2) \cdot e^{-\lambda \cdot W_1(m_1,2)} \right\} \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ \max\limits_{\substack{C_1(m_1,A_1) < \beta \\ m_1 = 1,2,3,\ldots}} \left\{ R_1(m_1,A_1) \cdot e^{-\lambda \cdot W_1(m_1,A_1)} \right\} \end{bmatrix} \qquad (3\text{-}34)$$

In the course of computing the sequence of function $f_K(\beta)$ for $K = 2,3,\ldots,n$, one can simultaneously obtain an optimal set of m_K and a_K (or all alternative solutions if the solution is not unique). This can be done as follows:

First compute the quantity $f_1(\beta)$ in (3-34) where in computing $f_1(\beta)$ for a given β, one must evaluate m_1 over all integers in the interval indicated for a given a_1.

Having computed $f_1(\beta)$ for each $\beta = 0,1,2,\ldots, C$, denote by $\hat{m}_1(\beta)$ and $\hat{a}_1(\beta)$, the value(s) of m_1 and a_1 for which

$$f_1(\beta) = R_1(\hat{m}_1(\beta),\hat{a}_1(\beta)) \cdot e^{-\lambda \cdot W_1(\hat{m}_1(\beta),\hat{a}_1(\beta))} \tag{3-35}$$

that is, $\hat{m}_1(\beta)$ and $\hat{a}_1(\beta)$ are the values of m_1 and a_1 which maximize the function $R_1(m_1,a_1)\ e^{-\lambda \cdot W_1(m_1,a_1)}$, when the amount β of the system cost is available for allocation to the unit, and m_1 and a_1 are constrained by $C_1(m_1,a_1) \leq \beta$.

The results of the computation for the first stage can now be arranged conveniently in a tabular form such as Table 1. If one wish to obtain all alternative solutions provided that the solution to (3-34) is not unique, one can so record all the alternative solutions in Table 1.

The right-most column gives the cumulative weight when using the optimal policy of $\hat{m}_1(\beta)$ and $\hat{a}_1(\beta)$ for a given β, and is determined by the relation

$$\bar{W}_K(\beta) = W_K(\hat{m}_K(\beta),\hat{a}_K(\beta)) + \bar{W}_{K-1}(\beta - C_K(\hat{m}_K(\beta),\hat{a}_K(\beta))) \tag{3-36}$$

Table 1. First Stage Policy Table

λ	β	$f_1(\beta)$	$\hat{m}_1(\beta)$	$\hat{a}_1(\beta)$	$\bar{W}_1(\beta)$
	0	$f_1(0)$	$\hat{m}_1(0)$	$\hat{a}_1(0)$	$\bar{W}_1(0)$
	1	$f_1(1)$	$\hat{m}_1(1)$	$\hat{a}_1(1)$	$\bar{W}_1(1)$
	2	$f_1(2)$	$\hat{m}_1(2)$	$\hat{a}_1(2)$	$\bar{W}_1(2)$
	3	$f_1(3)$	$\hat{m}_1(3)$	$\hat{a}_1(3)$	$\bar{W}_1(3)$
	.	.	.	.	.
	.	.	.	.	.
	.	.	.	.	.
	C	$f_1(C)$	$\hat{m}_1(C)$	$\hat{a}_1(C)$	$\bar{W}_1(C)$

for $K = 2,3,4,\ldots,n$. In particular for $K = 1$, one has

$$\bar{W}_1(\beta) = W_1(\hat{m}_1(\beta),\hat{a}_1(\beta)) \tag{3-27}$$

and for $K = n$

$$\bar{W}_n(C) = \sum_{K=1}^{n} W_K(m_K^*,a_K^*) \tag{3-28}$$

holds, where m_K^* and a_K^* are the optimal policy in the Kth stage, respectively. It is to be noted that the value of $\bar{W}_K(\beta)$ in the above policy table is the one continuously updated as the stage computations progress, which is similar to $f_K(\beta)$, but not a set of values for optimal policy of m_K and a_K. The quantity $\bar{W}_n(C)$ in (3-38) is the value of $\sum W_K(m_K^*, a_K^*)$ used to determine the next value of Lagrange multiplier λ.

Once the $\bar{W}_K(\beta)$ table is constructed at each stage, it may be preferable for a small-size digital computer not to store all intermediate computational results including the optimal policy at each stage except the $\bar{W}_K(\beta)$ table while searching for a correct value of Lagrange multiplier λ which yields an optimal solution to the problem. Upon determining the correct λ, the computation can be done again, storing this time all optimal policy tables.

Having obtained $f_1(\beta)$, one now proceeds to compute $f_2(\beta)$ for every $\beta = 0,1,2,\ldots,C$, using the recurrence relation given in (3-33). Thus

$$f_2(\beta) = \max \begin{bmatrix} \max\limits_{\substack{C_2(m_2,1) \le \beta \\ m_2 = 1,2,3,\ldots}} \left\{ R_2(m_2,1) \cdot e^{-\lambda \cdot W_2(m_2,1)} \cdot f_1(\beta - C_2(m_2,1)) \right\} \\ \max\limits_{\substack{C_2(m_2,2) \le \beta \\ m_2 = 1,2,3,\ldots}} \left\{ R_2(m_2,2) \cdot e^{-\lambda \cdot W_2(m_2,2)} \cdot f_1(\beta - C_2(m_2,2)) \right\} \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ \max\limits_{\substack{C_2(m_2,A_2) \le \beta \\ m_2 = 1,2,3,\ldots}} \left\{ R_2(m_2,A_2) \cdot e^{-\lambda \cdot W_2(m_2,A_2)} \cdot f_1(\beta - C_2(m_2,A_2)) \right\} \end{bmatrix} \tag{3-39}$$

In carrying out the maximization in (3-39), one lets m_2 assume all integers in the interval defined, and computes, for a given a_2, the quantity

$$R_2(m_2,a_2) \cdot e^{-\lambda \cdot W_2(m_2,a_2)} \cdot f_1(\beta - C_2(m_2,a_2)) \qquad (3\text{-}40)$$

for each value of m_2, and selects the largest of these values so obtained. Repeating the same maximization for every value of $a_2 = 1,2,3,\ldots,A_2$, and obtains $f_2(\beta)$ which is the largest of these A_2 values for a given β. Simultaneously, one also determines the values of $\hat{m}_2(\beta)$ and $\hat{a}_2(\beta)$ which are the maximizing values of $f_2(\beta)$, respectively.

It is to be noted that to obtain $f_2(\beta)$, one needs to know first $f_1(\beta - C_2(m_2,a_2))$ for all integral values that m_2 takes on. Note also that the quantity $\beta - C_2(m_2,a_2)$, which appears as the argument of f_1 is also a nonnegative integer since it was originally assumed that the function $C_K(m_K,a_K)$ assumes only integral values for $K = 1,2,3,\ldots,n$. Since one has already evaluated $f_1(\beta)$ for every integer $\beta = 0,1,2,\ldots,C$, one can readily obtain the value of $f_1(\beta - C_2(m_2,a_2))$ from the Table 1 by reading the value of $f_1(\beta)$ corresponding to the number $\beta - C_2(m_2,a_2)$ in β column. One is thus able to construct the table of the same form as Table 1 for $f_2(\beta)$, $\hat{m}_2(\beta)$, $\hat{a}_2(\beta)$, and $\bar{W}_2(\beta)$.

The next step is to compute $f_3(\beta)$ for each $\beta = 0,1,2,\ldots,C$, by means of the recurrence relation in (3-33) for $K = 3$, and also using the tables for $f_2(\beta)$. The tables of $f_3(\beta)$, $\hat{m}_3(\beta)$, $\hat{a}_3(\beta)$ and $\bar{W}_3(\beta)$ can then be constructed at this point. This procedure is repeated until one is finally ready to evaluate $f_n(\beta)$. At this point one needs only to evaluate $f_n(C)$ unless one is particularly interested in a sensitivity analysis to see how the quantity of C available influences the optimal solution for a given reliability allocation problem. The value $\hat{m}_n(C) = m_n^*$ and $\hat{a}_n(C) = a_n^*$, that is, the values of m_n and a_n which yield $f_n(C)$ is then simply m_n^* and a_n^* .

Having obtained m_n^* and a_n^*, the remaining n-1 variables must satisfy the relation

$$\sum_{j=1}^{n-1} C_j(m_j,a_j) \leq C - C_n(m_n^*,a_n^*) \tag{3-41}$$

in maximizing

$$\prod_{j=1}^{n-1} R_j(m_j,a_j)\, e^{-\lambda \cdot W_j(m_j,a_j)} .$$

But this maximum is simply $f_{n-1}(C - C_n(m_n^*,a_n^*))$, and the values of m_{n-1} and a_{n-1} which yields $f_{n-1}(C - C_n(m_n^*,a_n^*))$ are $\hat{m}_{n-1}(C - C_n(m_n^*,a_n^*))$, and $\hat{a}_{n-1}(C - C_n(m_n^*,a_n^*))$, respectively.

Therefore

$$m_{n-1}^* = \hat{m}_{n-1}(C - C_n(m_n^*,a_n^*)) \tag{3-42}$$

$$a_{n-1}^* = \hat{a}_{n-1}(C - C_n(m_n^*,a_n^*))$$

which can be read directly from the table of $\hat{m}_{n-1}(\beta)$ and $\hat{a}_{n-1}(\beta)$ for

$\beta = C - C_n(m_n^*, a_n^*)$. In the same way, one obtains

$$m_{n-2}^* = \hat{m}_{n-2}\ (C - C_n(m_n^*, a_n^*) - C_{n-1}(m_{n-1}^*, a_{n-1}^*))$$

$$a_{n-2}^* = \hat{a}_{n-2}\ (C - C_n(m_n^*, a_n^*) - C_{n-1}(m_{n-1}^*, a_{n-1}^*))$$

and in general

$$m_{n-K}^* = \hat{m}_{n-K}\ (C - \sum_{t=0}^{K-1} C_{n-t}\ (m_{n-t}^*, a_{n-t}^*)) \qquad (3\text{-}43)$$

$$a_{n-K}^* = \hat{a}_{n-K}\ (C - \sum_{t=0}^{K-1} C_{n-t}\ (m_{n-t}^*, a_{n-t}^*))$$

Thus one has obtained all the m_K^* and a_K^* for $K = 1,2,3,\ldots,n$.

Maximum system reliability achieved by using the optimal policy is then simply

$$R(m_1^*, m_2^*, \ldots, m_n^*;\ \ a_1^*, a_2^*, \ldots, a_n^*) = f_n(C) \cdot e^{\lambda \cdot \bar{W}_n(C)} \qquad (3\text{-}44)$$

with respect to a given Lagrange multiplier λ, since $\bar{W}_n(C) = \sum_{K=1}^{n} W_K(m_K^*, a_K^*)$.

The computational procedure described above makes it possible to determine the optimal policy associated with the given n-serial unit system for a given value of the Lagrange multiplier λ. As has been noted, there is no precise prescription as to what value of λ will produce a desired optimal solution to the problem, satisfying the relation

$$\sum_{K=1}^{n} W_K(m^*(\lambda),a^*(\lambda)) = \bar{W}_n(C) = W \tag{3-45}$$

If the first guess is λ_o, and $W_n(C;\ \lambda_o) > W$, then one should increase for the next trial so as to decrease the total system weight. If one can find a λ such that the eliminated weight constraint holds as an equality, then from what has been proved in the preceding section, the solution so obtained is indeed an optimal solution to the problem.

In many cases, linear interpolation using the two previous values of λ, say, λ_1, and λ_2 and the resulting values of $\bar{W}_n(C;\ \lambda_1)$ and $\bar{W}_n(C;\ \lambda_2)$ will provide a good estimate for the third and subsequent values of λ which should yield $\bar{W}_n(C;\ \lambda) = W$. The linear interpolation formula for determining a next value, λ_3, is given by

$$\lambda_3 = \lambda_1 + \frac{\lambda_2 - \lambda_1}{\bar{W}_n(C;\lambda_2) - \bar{W}_n(C;\lambda_1)} (W - \bar{W}_n(C;\lambda_1)) \tag{3-46}$$

Usually one can converge on the correct value of λ in five or six iterations, and thus solve the problem.

The computational procedure above shows that in computing the elements of the sequence $\{f_K(\beta)\}$, $K = 1,2,3,\ldots,n-1$, the value of function must be tabulated for all $\beta = 0,1,2,\ldots,C$. Unless one is particularly interested in a sensitivity study of the optimal solution, all that is ultimately required is $f_n(C)$. However, to obtain $f_n(C)$, one must carry out a maximization over m_n and a_n and compute $f_{n-1}(C - C_n(m_n,a_n))$ for all m_n and a_n such that $C_n(m_n,a_n) \le C$. The value of the argument β, for which one needs $f_{n-1}(\beta)$ will, of course, depends on $C_n(m_n,a_n)$.

If the cost of the nth functional unit for all alternatives is unity, one must know $f_{n-1}(\beta)$ for all $\beta = 0,1,2,\ldots,C-1$. On the other hand, if the values of C and $C_n(m_n,a_n)$ take on only even numbers, one needs only to evaluate the function for even numbers of β. For each β for which $f_{n-1}(\beta)$ is needed, one also needs $f_{n-2}(\beta - C_{n-1}(m_{n-1},a_{n-1}))$ for all nonnegative integers m_{n-1}, and a_{n-1} such that $C_{n-1}(m_{n-1},a_{n-1}) \leq \beta$. It then becomes evident that for a large number of variables it is virtually impossible to trace back to determine precisely what value of β one needs to tabulate $f_K(\beta)$ at each stage. Thus, instead of making this detailed analysis, it is much more preferable to tabulate $f_K(\beta)$ for each value of β from zero to C. In this way one can always guarantee having all necessary values of $f_K(\beta)$ in the succeeding computations.

Since at each stage one needs to evaluate $f_K(\beta)$ for $\beta = 0,1,2,\ldots,C$, the original restriction that the function $C_K(m_K,a_K)$ must take on integral numbers will ensure that the arguments of the $f_K(\beta)$ are always integers. This requirement will make it easy to determine the set of values for which $f_K(\beta)$ needs to be evaluated. This restriction actually does not present any problem in practice, since to an arbitrary degree of accuracy in measuring the cost, this can always be accomplished without difficulty. The restriction merely presents a problem of properly selecting the monetary denomination in measuring the cost so as to ensure the integral values.

Discussion on the Computational Procedures

It is now evident from the foregoing discussions that the dynamic programming approach to solving the reliability allocation problem is an efficient computational algorithm and always yields an optimal solution to the problem. Once the unit reliability function $R_K(m_K, a_K)$, and the constraint function $C_K(m_K, a_K)$ are specified, the quantity β of the cost available for allocation to the K stages is the only parameter which influences the decision variables m_K and a_K at each stage. Thus the quantity β describes the state of the system and is often called state variable.

The dynamic programming approach used here not only solves the system reliability allocation problem for $1,2,\ldots,n$ stages, but at the same time, determines for each $K < n$ an optimal solution for every value of β up to the given amount C of the cost ultimately available for all n stages. This basic feature of the computational procedure clearly indicates the important fact that addition of one more unit to the existing system or any change in the value of the parameters in the problem does not present any problem as far as the computational effort is concerned. The nature of the dynamic programming approach is such that the optimal solution for $n + 1$ stages is obtained from the existing optimal solution for n stages by simply adding the $(n + 1)$ st stage and making use of the existing solution for n stages.

It is rather common that in the early phases of the system design, the system designer is not only concerned with the optimal allocations of the system reliability requirement for any predetermined level of system cost or system weight parameter, but also allowing the parameter to vary over a critical range of values, and then observing how the optimal solution is affected by these changes in the parameter. This aspect of the problem is intimately connected with the concept of a sensitivity analysis.

Unless one has a fairly simple and explicit analytical solution to the problem, it is practically impossible to obtain this sort of information through the conventional approach. In this respect, the dynamic programming approach makes it possible without adding much computational effort to supply a variety of important information on the sensitivity analysis. This is because the optimal solution by the dynamic programming approach is always given as a function of the basic system parameter C and n, and consequently a large class of sensitivity analysis information automatically accompanies the optimal solution in the course of computation.

It might be entirely possible to set up the dynamic programming formulation of the present problem even if one had not introduced at the outset the Lagrange multiplier which results in a removal of the weight constraint, and reduces by one the dimensionality of the problem. However, it is to be noted in this case that the state functions $f_K(\beta,\mu)$ are the maximum reliability achieved from first K stages when β and μ units of system cost and system weight, respectively, are available for allocation to these stages, and are now functions of two arguments.

For the one state parameter problem which has been discussed so far, it is, in general, necessary to compute and tabulate $f_K(\beta)$

and associated $\hat{m}_K(\beta)$ and $\hat{a}_K(\beta)$ for every $\beta = 0,1,2,\ldots C$. Exactly the same argument applies that for the allocation problem having two state parameters, it may be necessary to tabulate $f_K(\beta,\mu)$ and $\hat{m}_K(\beta,\mu)$ and $\hat{a}_K(\beta,\mu)$ for every possible combination of β and μ, where β now ranges over $0,1,2,\ldots,C$, and μ also ranges over $0,1,2,\ldots,W$.

Thus, if it is possible for both β and μ to assume 500 different values, then the total of 250,000 maximization computations must be carried out at each stage, whereas only 500 such computations are required for the same situation when one can reduce this problem to the one-dimensional problem. The effort required for computing and tabulating $f_K(\beta,\mu)$ may easily be at least 500 times as great as the corresponding case of the one-dimensional problem.

Any attempt to compute in a straightforward way and store these values of $f_K(\beta,\mu)$ in the computer may be prohibited from the huge memory requirements even in the moderate-size digital computer. Sometimes, just a single table of $f_K(\beta,\mu)$ could easily exceed the memory capacity of the high-speed computer, thus requiring the use of magnetic tape which undoubtedly slows down the computational speed of the computer.

Therefore, it can be safely asserted that reduction of the dimensionality of the problem by introducing the Lagrange multiplier indeed requires considerably less work than the straightforward formulation of the problem. One could try 200 or more values of λ before the computational effort might be comparable with the straightforward method. Furthermore, as has been noted, each value of λ which has been selected as a trial value will yield an optimal solution to the problem with a specified level of system weight.

It may be necessary to adjust the value of λ by trial and error to achieve any given level of weight constraint stated in advance, but each computation for a given value of λ is by no means wasteful, and these computations rather furnish important information on the sensitivity of the optimal solution for a range of the weight constraint.

Some Numerical Results

In the preceding section the actual steps of a computational solution to the system reliability allocation problem are described in detail. In order to illustrate this computational procedure, a following hypothetical system reliability allocation problem is considered. The system under consideration consists of 14 functional units which are connected in series, and failure probability of each unit in the system is assumed to be independent.

If m_K of the identical units are used as a parallel redundancy in the Kth unit, the reliability of the Kth unit is given by

$$R_K = [1 - (1 - R_{Kj})]^{m_K}$$

where R_{Kj} denotes the unit reliability of the Kth unit using jth design alternative. It is further assumed that a unit of the Kth unit using Jth design alternative costs C_{Kj} and weighs W_{Kj}.

Given the total amount of 130 units of system cost and 170 units of system weight constraints, the reliability allocation problem is to find the optimum number of redundancy to use and the optimum design alternative at each stage which will result in the greatest system reliability while keeping the total cost and weight less than the given amount. The necessary input data for this problem are summarized in Table 2.

The detailed procedural steps for computing the optimum solution to the problem are displayed in the flow chart shown in the next section. This flow chart is used for the computer programming of the problem. Table 3 is a reproduction of the section of the computer output table showing the values of the functions $f_{14}(C)$, $\hat{m}_{14}(C)$, $\hat{a}_{14}(C)$ and $\bar{W}_{14}(C)$.

Table 4, then, presents the optimal solution to the problem. The value of Lagrange multiplier λ which yields the required system weight of 170 units and thus gives the optimal solution to the reliability allocation problem is 0.0012.

Total system cost associated with the optimal policy is 119, and the total system weight is 170. System reliability associated with the optimal policy found in Table 4 is then

$$R = f_{14}(130) \cdot e^{\lambda \cdot \bar{W}_{14}(130)}$$

$$= (0.7910) \cdot e^{0.0012 \times 170}$$

$$= 0.9700$$

In the course of searching for a correct value of the Lagrange multiplier λ, which should yield a system weight of 170 units, several

Table 2. Numerical Input Data for a Reliability Allocation Problem

Functional Unit (I)	DESIGN ALTERNATIVE (J)											
	1			2			3			4		
	R	C	W	R	C	W	R	C	W	R	C	W
1	0.90	1	3	0.93	1	4	0.91	2	2	0.95	2	5
2	0.95	2	8	0.94	1	10	0.93	1	9	*	*	*
3	0.85	2	7	0.90	3	5	0.87	1	6	0.92	4	4
4	0.83	3	5	0.87	4	6	0.85	5	4	*	*	*
5	0.94	2	4	0.93	2	3	0.95	3	5	*	*	*
6	0.99	3	5	0.98	3	4	0.97	2	5	0.96	2	4
7	0.91	4	7	0.92	4	8	0.94	5	9	*	*	*
8	0.81	3	4	0.90	5	7	0.91	6	6	*	*	*
9	0.97	2	8	0.99	3	9	0.96	4	7	0.91	3	8
10	0.83	4	6	0.85	4	5	0.90	5	6	*	*	*
11	0.94	3	5	0.95	4	6	0.96	5	6	*	*	*
12	0.79	2	4	0.82	3	5	0.85	4	6	0.90	5	7
13	0.98	2	5	0.99	3	5	0.97	2	6	*	*	*
14	0.90	4	6	0.92	4	7	0.95	5	6	0.99	6	9

* Denotes no design alternative available for this unit.

Table 3. Values of the Functions

$f_{14}(C)$, $\hat{m}_{14}(C)$, $\hat{a}_{14}(C)$, and $\bar{W}_{14}(C)$ for $\lambda = 0.0012$

Cost C	$f_{14}(C)$	$\hat{m}_{14}(C)$	$\hat{a}_{14}(C)$	$\bar{W}_{14}(C)$
90	0.7771	1	4	157
91	0.7785	1	4	154
92	0.7791	1	4	161
93	0.7796	2	3	153
94	0.7802	2	3	160
95	0.7816	2	3	157
96	0.7822	2	3	164
97	0.7826	2	3	157
98	0.7831	2	3	164
99	0.7836	2	3	150
100	0.7842	2	3	157
101	0.7856	2	3	154
102	0.7862	2	3	161
103	0.7866	2	3	154
104	0.7872	2	3	161
105	0.7874	2	3	157
106	0.7879	2	3	164
107	0.7879	2	3	164
108	0.7885	2	3	159
109	0.7886	2	3	168
110	0.7893	2	3	162
111	0.7893	2	3	162
112	0.7895	2	3	162
113	0.7899	2	3	166
114	0.7902	2	3	166
115	0.7902	2	3	166
116	0.7904	2	3	166
117	0.7908	2	3	170
118	0.7908	2	3	170
119	0.7910	2	3	170
120	0.7910	2	3	170
121	0.7910	2	3	170
122	0.7910	2	3	170
123	0.7910	2	3	170
124	0.7910	2	3	170
125	0.7910	2	3	170
126	0.7910	2	3	170
127	0.7910	2	3	170
128	0.7910	2	3	170
129	0.7910	2	3	170
130	0.7910	2	3	170

Table 4. Solution of the Reliability Allocation Problem

Unit Number K	Number of Units Required m_K^*	Design Alternative a_K^*
1	3	3
2	2	1
3	3	4
4	3	3
5	3	2
6	2	2
7	2	1
8	4	1
9	2	3
10	3	2
11	2	1
12	4	1
13	2	2
14	2	3

Table 5. Maximum System Reliabilities for Various Values of Lagrange Multipliers

Lagrange Multiplier λ	Maximum System Reliability R	System Cost C	System Weight W
0.0004	0.9864	130	191
0.0005	0.9864	130	191
0.0006	0.9815	130	182
0.0007	0.9815	130	182
0.0008	0.9792	127	179
0.0009	0.9792	127	179
0.0010	0.9764	125	176
0.0011	0.9744	121	174
0.0012	0.9700	119	170
0.0013	0.9700	119	170
0.0014	0.9646	116	166
0.0015	0.9589	112	162
0.0016	0.9546	110	159

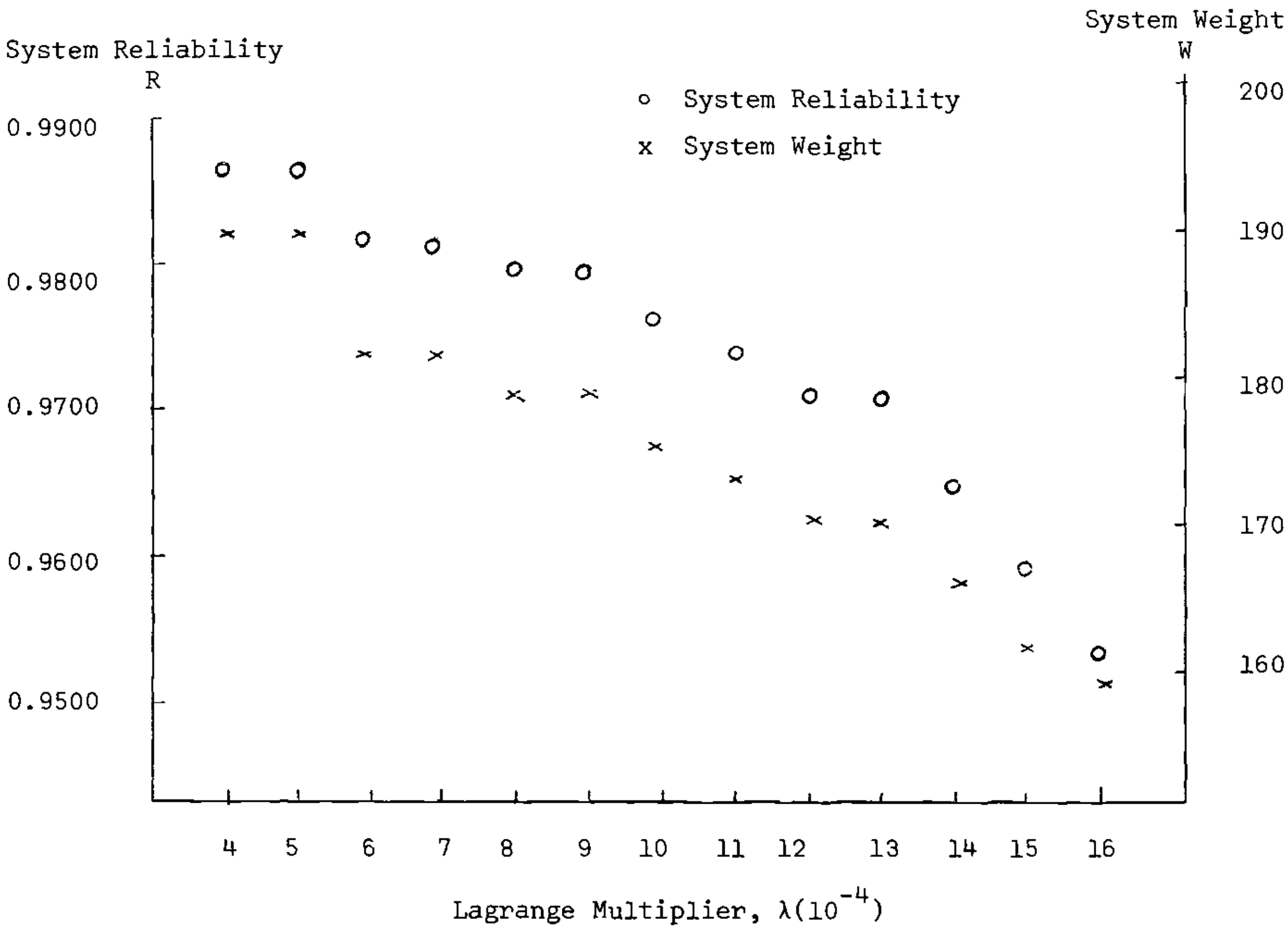


Figure 4. System Reliability and Weight as a Function of Lagrange Multiplier

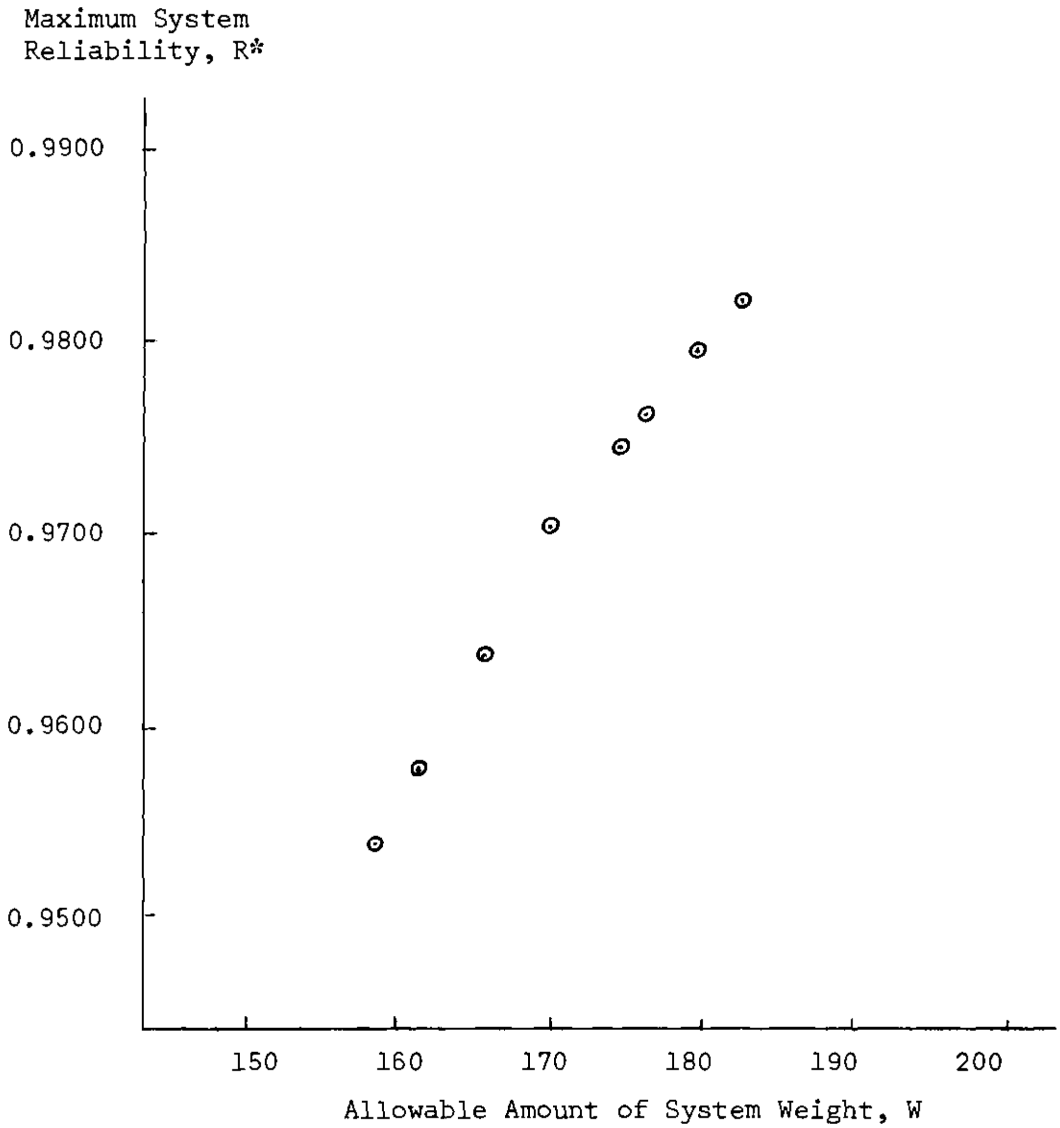


Figure 5. Maximum System Reliability as a Function of System Weight

values of λ have been tried, and the resulting summary of the solutions is shown in Table 5. As has been verified earlier, it is also seen from the data in Table 5 that maximum system reliability and total system weight will decrease as the value of Lagrange multiplier increases. In Figure 4, maximum system reliability and total system weight are plotted as a function of the Lagrange multiplier to demonstrate this monotonicity in λ. Figure 5 shows how the maximum system reliability depends upon the allowable amount of system weight, and thus gives a fairly good picture on the sensitivity of the optimal solution.

As a second numerical example, the following hypothetical system consisting of 20 functional units is considered. Again it is assumed that maximum allowable amount of system cost is 130 units, and the system weight is 233 units. The input data used is presented in Table 6.

The optimal solution for this reliability allocation problem is then shown in Table 7. The value of Lagrange multiplier associated with the optimal solution is 0.001. The total system cost and total system weight associated with the optimum policy found above is 130 and 233 units, respectively. Maximum system reliability is then

$$R = f_{20}(130) \cdot e^{\lambda \cdot \bar{W}_{20}(130)}$$

$$= (0,7511) \cdot e^{0.001 \times 233}$$

$$= 0.9481$$

Table 6. Numerical Input Data for a Hypothetical System

Functional Unit (I)	DESIGN ALTERNATIVE (J) 1 R	1 C	1 W	2 R	2 C	2 W	3 R	3 C	3 W	4 R	4 C	4 W
1	0.90	1	3	0.93	1	4	0.91	2	2	0.95	2	5
2	0.95	2	8	0.94	1	10	0.93	1	9	*	*	*
3	0.85	2	7	0.90	3	5	0.87	1	6	0.92	4	4
4	0.83	3	5	0.87	4	6	0.85	5	4	*	*	*
5	0.94	2	4	0.93	2	3	0.95	3	5	*	*	*
6	0.99	3	5	0.98	3	4	0.97	2	5	0.96	2	4
7	0.91	4	7	0.92	4	8	0.94	5	9	*	*	*
8	0.81	3	4	0.90	5	7	0.91	6	6	*	*	*
9	0.97	2	8	0.99	3	9	0.96	4	7	0.91	3	8
10	0.83	4	6	0.85	4	5	0.90	5	6	*	*	*
11	0.94	3	5	0.95	4	6	0.96	5	6	*	*	*
12	0.79	2	4	0.82	3	5	0.85	4	6	0.90	5	7
13	0.98	2	5	0.99	3	5	0.97	2	6	*	*	*
14	0.90	4	6	0.92	4	7	0.95	5	6	0.99	6	9
15	0.91	1	5	0.92	2	4	0.94	3	6	*	*	*
16	0.90	2	4	0.91	3	5	0.92	4	4	*	*	*
17	0.87	1	2	0.88	2	4	0.90	2	5	0.91	3	6
18	0.85	2	4	0.87	2	5	0.94	3	7	*	*	*
19	0.92	3	5	0.97	4	6	0.98	4	8	*	*	*
20	0.90	1	4	0.92	2	5	0.94	2	6	0.95	3	7

* Denotes no fourth design alternative available.

Table 7. Optimal Solution of the Problem

Functional Unit Number K	Number of Units To Use m_K^*	Design Alternative a_K^*
1	3	3
2	2	1
3	2	4
4	3	1
5	3	2
6	2	4
7	2	1
8	3	1
9	2	1
10	3	2
11	2	1
12	4	1
13	2	1
14	2	3
15	3	2
16	3	1
17	3	1
18	3	1
19	2	2
20	3	1

The computer time expended in the computation of the above problem is seven minutes, whereas the time for the previous problem with 14-unit system is five minutes for two values of Lagrange multiplier. The ALGOL Program used in the computation of the above problem is included in the Appendix.

Concluding Remarks

With the application of the functional allocation method, system designers can obtain a first set of numerical reliability requirements for the major functional equipment levels. However, the application of this method will only yield early decisions for the reliability requirements at these levels, and must therefore be reviewed and modified early in the design stages as soon as the detailed feasibility study discloses the discrepancies between allocated requirement and improvement feasibility. The most interesting case arises when there are several design alternatives available for achieving the needed level of reliability requirements including the conventional approach of employing redundancy for the weak units.

In this chapter, a reliability allocation method has been developed to solve the problem of supplementary reliability allocations within the functional equipment groups when (a) there are a number of design alternatives for achieving the needed level of reliability requirements, and (b) there also exist the constraints on system costs and weights in a given system. The reliability allocation problem formulated in Equations (3-1) and (3-2), then, is to select the design alternative and the level of redundancy at each subsystem in such a way as to maximize the overall system reliability.

The dynamic programming approach to this multi-stage allocation problem is believed to be an efficient one since the computational scheme involved is simple, practical, easily implemented, and yet yields the exact solution to the problem. The method described in this chapter is also applicable to a class of sequential decision problems involving two discrete decision variables with a set of constraints.

The modified formulation of the problem shown in Equations (3-3) and (3-4) in an attempt to reduce the dimensionality of the problem will reduce the amount of computational effort considerably, and this reduction is especially impressive as the number of stages in the system increases. Whenever the modified allocation problem is solved for a given set of the Lagrange multipliers, a by-product of the calculation of the optimal policy is a set of state values (the values of an eliminated constraint function in the optimal policy) that permit the sensitivity analysis in the given circumstances. In most cases of the system reliability allocations, this sensitivity analysis is more interesting and useful than solving the problem for a particular value of the constraint.

Flow Chart for the Computational Procedures

Description of Input Data

N = number of functional units in the system

$\bar{C}$ = maximum allowable amount of money

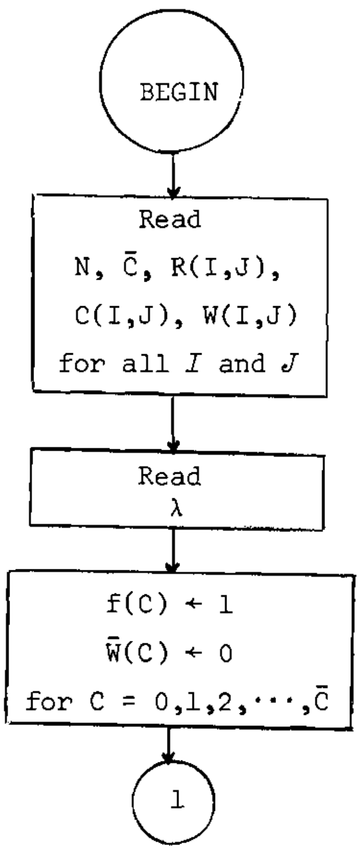
λ = Lagrange multiplier

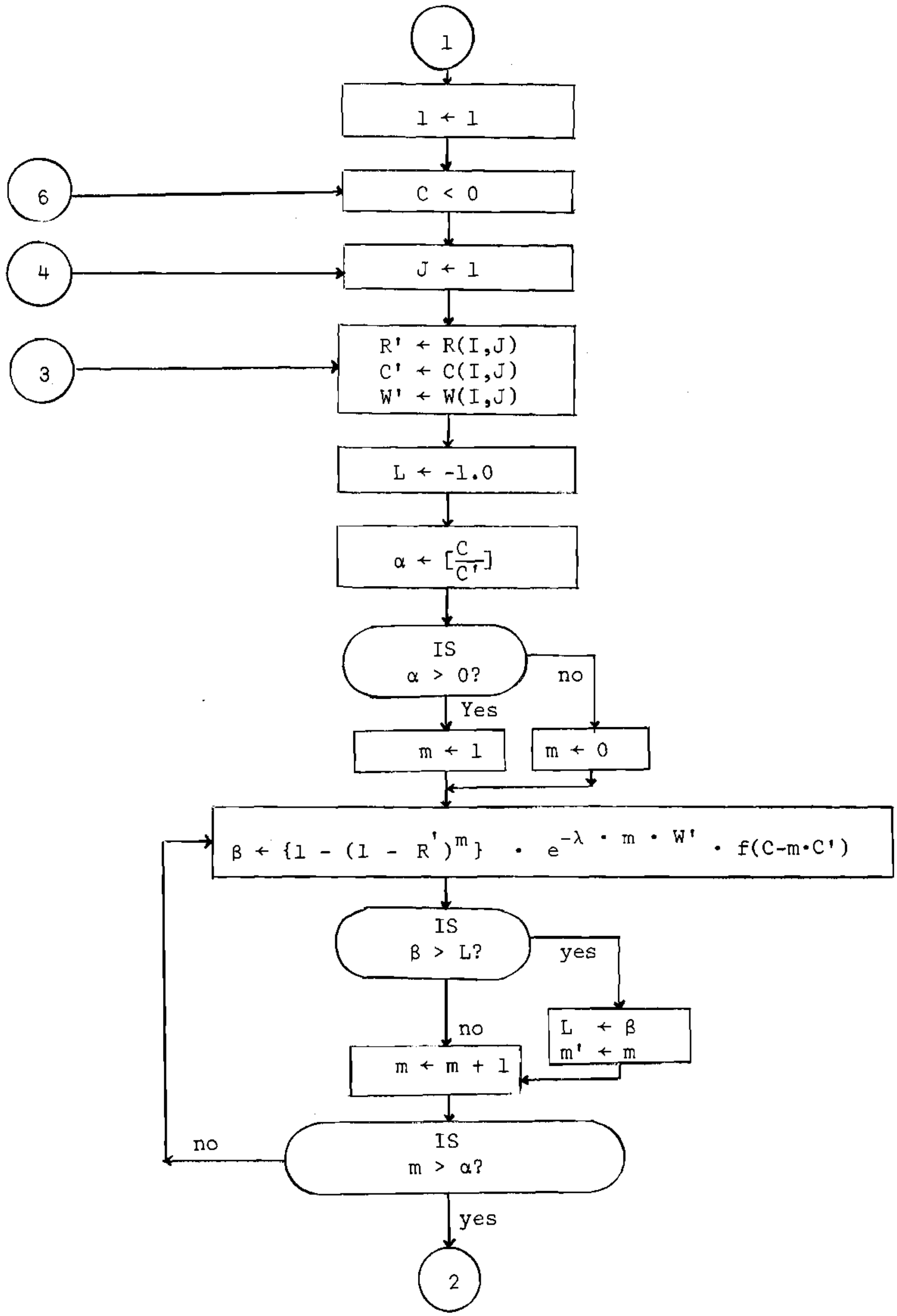
A(I) = number of design alternatives available for the *I*th functional unit

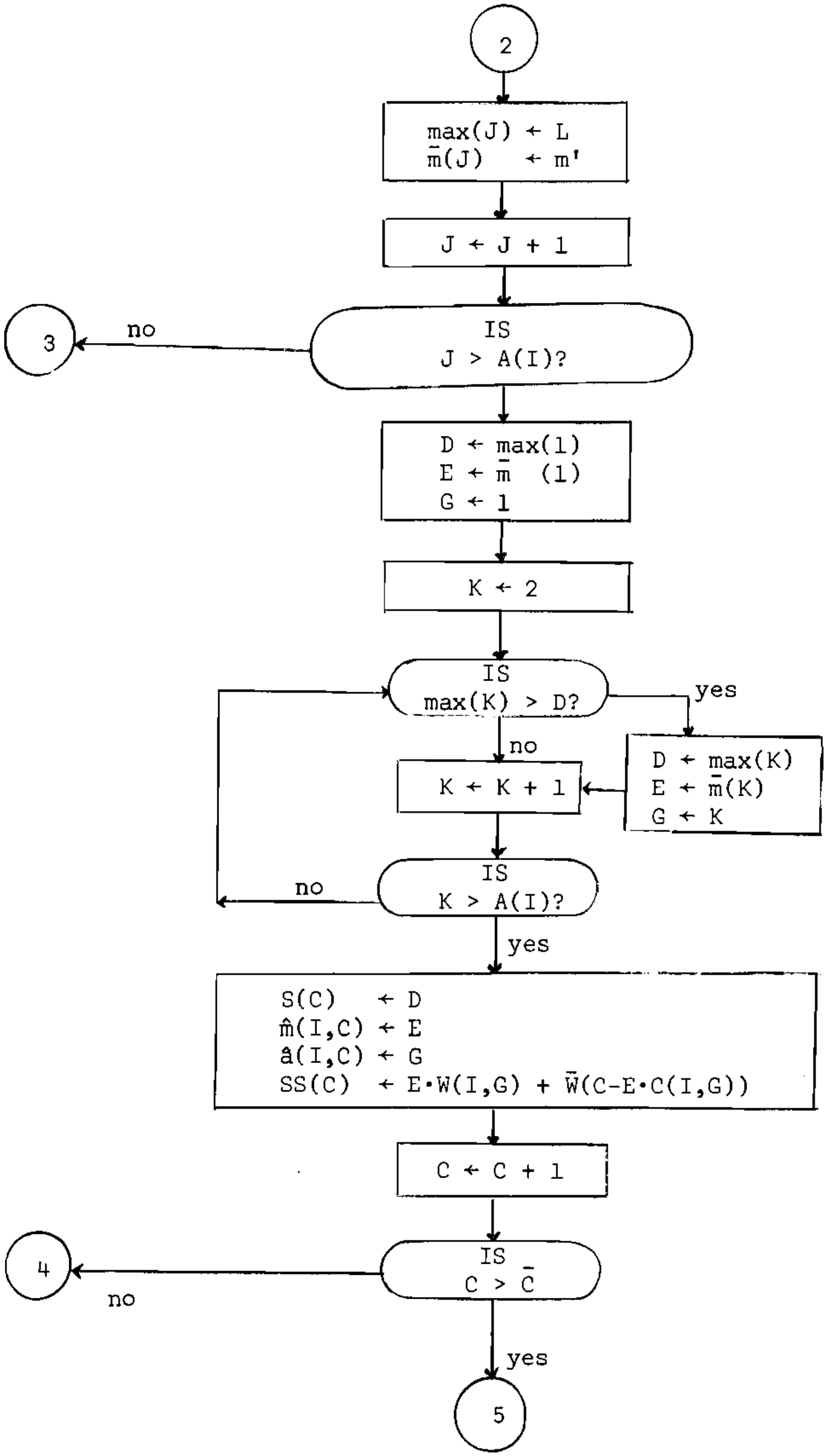
R(I,J) = unit reliability of the *I*th unit, using the *J*th design alternative

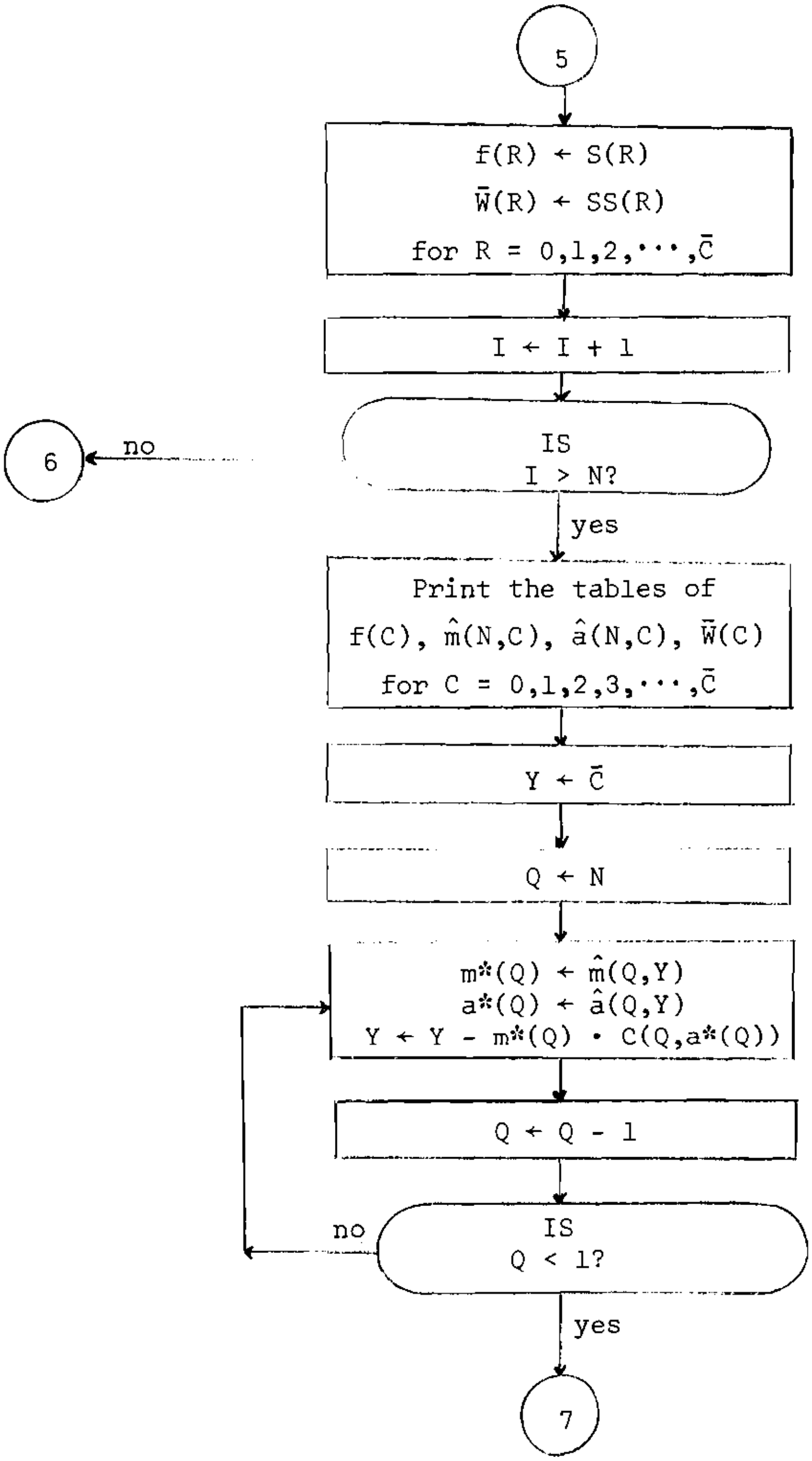
C(I,J) = unit cost of the *I*th unit, using the *J*th design alternative

W(I,J) = unit weight of the *I*th unit, using the *J*th design alternative

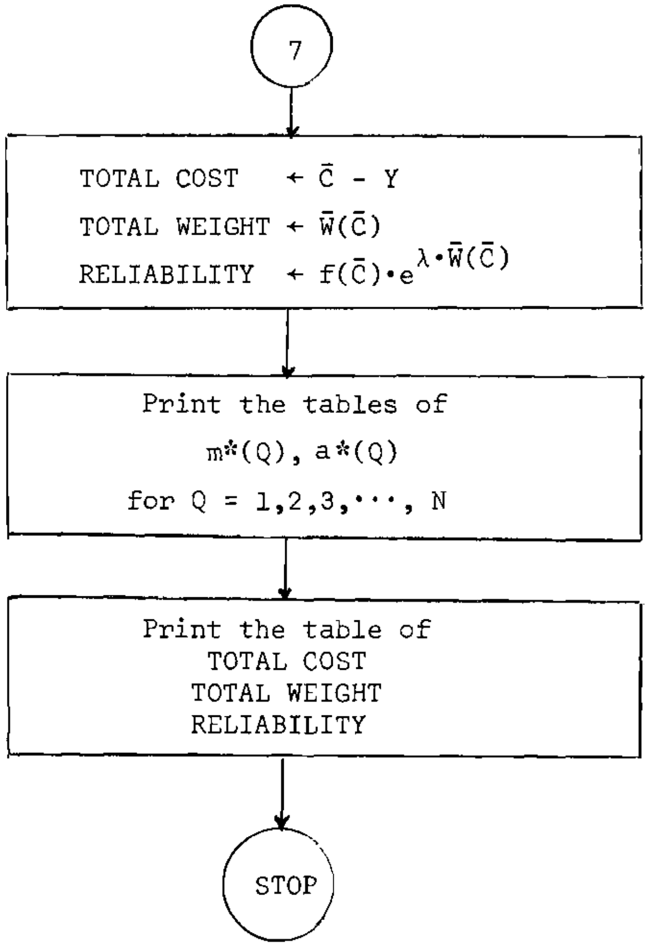



1
I ← 1
6
C < 0
4
J ← 1
3
R' ← R(I,J)
C' ← C(I,J)
W' ← W(I,J)
L ← -1.0
α ← [C/C']
IS
α > 0?
no
Yes
m ← 1
m ← 0
β ← {1 - (1 - R')^m} · e^(-λ · m · W') · f(C-m·C')
IS
β > L?
yes
L ← β
m' ← m
no
m ← m + 1
no
IS
m > α?
yes
2


2
max(J) ← L
m̄(J) ← m'
J ← J + 1
IS
J > A(I)?
no
3
D ← max(1)
E ← m̄ (1)
G ← 1
K ← 2
IS
max(K) > D?
yes
D ← max(K)
E ← m̄(K)
G ← K
no
K ← K + 1
IS
K > A(I)?
no
yes
S(C) ← D
m̂(I,C) ← E
â(I,C) ← G
SS(C) ← E·W(I,G) + W̄(C-E·C(I,G))
C ← C + 1
IS
C > C̄
no
4
yes
5



5
f(R) ← S(R)
W̄(R) ← SS(R)
for R = 0,1,2,···,C̄
I ← I + 1
IS
I > N?
no
6
yes
Print the tables of
f(C), m̂(N,C), â(N,C), W̄(C)
for C = 0,1,2,3,···,C̄
Y ← C̄
Q ← N
m*(Q) ← m̂(Q,Y)
a*(Q) ← â(Q,Y)
Y ← Y - m*(Q) · C(Q,a*(Q))
Q ← Q - 1
IS
Q < 1?
no
yes
7


7
TOTAL COST ← $\bar{C}$ - Y
TOTAL WEIGHT ← $\bar{W}(\bar{C})$
RELIABILITY ← $f(\bar{C}) \cdot e^{\lambda \cdot \bar{W}(\bar{C})}$
Print the tables of
m*(Q), a*(Q)
for Q = 1,2,3,···, N
Print the table of
TOTAL COST
TOTAL WEIGHT
RELIABILITY
STOP

CHAPTER IV

RELIABILITY ALLOCATION BASED ON THE SYSTEM COSTS

Preliminary Considerations

The functional allocation method described in Chapter II permits numerical allocations of system reliability requirements at the functional equipment level. Within the restrictions of system costs and other performance requirements, the reliability allocations so obtained can be used in solving a reliability problem associated with the development of an effective system for the ultimate user. The extent to which the method will serve this purpose largely depends upon the manner in which it is interpreted and implemented in a given system.

The functional allocation method is largely concerned with the state-of-the-art and functional role of the major system elements judged from previous experiences and their projected use in a future system application. Generally, these are not the only dominant factors influencing the initial allocations of the system reliability. The concepts of cost and system effectiveness have become increasingly important in recent years. There is a growing awareness that reliability is not an entity which can be isolated from the total system.

A prime objective in the design and construction of any system is that it be capable of performing its intended function at the lowest possible system cost. In this context, introduction of a system cost concept in the reliability allocation process is essential for such purposes as (a) comparing alternative design policies, (b) minimizing relevant system costs which vary with the level of allocated reliability, and (c) determining the number of subsystems that are required in a given system.

Within any system development program, individuals responsible for allocating system reliability are faced with a continual series of compromises between reliability and system costs. By building a system with higher reliability, initial investment costs obviously increase. On the other hand, this addition of reliability at the expense of higher initial investment costs should lower the support costs over the useful life of a system. Thus, this type of trade-off between two costs involved is often required at the major functional equipment and support system levels to select the best design alternative early in the development program.

The conceptual approach of an allocation model based on the system costs would be to apportion the allowable system failure rate (corresponding to the specified system reliabilility requirement) among the constituent subsystems in such a way as to minimize the total system costs. In this chapter, it is intended to develop a reliability allocation method based on the relevant system costs which is consistent and applicable to various types of systems, rather than a method which is detailed enough for use in the particular system. The emphasis is therefore on the development of a conceptual framework of the allocation procedure.

The Relevant Cost Elements Involved in the Allocation

In order to translate the reliability allocation problem into a mathematical problem of minimizing a total system cost function, it is necessary first to identify the separate elements of system costs which vary with changes in the allocated subsystem failure rates, and then to derive a mathematical form that is both sufficiently flexible to approximate a wide range of complex cost relationships, and sufficiently simple to allow a mathematical solution to the problem.

Ideally, a specific cost function which can be used to determine a numerical cost value for a given level of subsystem failure rate should be and can be derived using conventional statistical techniques. However, the number of observations available from historical cost data in many practical situations is seldom sufficiently large to apply statistical methods such as regression. Even though the availability of such historical cost data is severely limited in many cases, empirical evidence supported by engineering experiences should at least be the basis for the derivation of a cost function which gives reasonable and comparable results.

Since the operation of a proposed system, after it has been built, will continue over relatively long periods of time, it is considered important to distinguish between the way in which relevant system costs are incurred for one-time initial investments and the requirements for continuing regular annual operating expenses. This distinction between one-time initial investment costs and annual recurring costs is useful not only for determining the extent to which there are differences in the total amounts of money or resources but also for the purpose of providing a basis for judgment so that one can clearly identify those elements of cost which must be met each time period if the system is to continue in operation, as distinguished from one-time investments.

One-time initial investment costs include such outlays as subsystem Research and Development, engineering, hardware fabrication and installation, and subsystem test costs. Subsystem Research and Development costs refer to the outlay required for the basic and applied research including the costs of R & D facilities and prototype development. The costs for engineering and hardware fabrication and installation represent primarily the costs of engineering, fabrication, tooling and installation. Subsystem test costs consist of such outlays as required for performing the subsystem reliability tests, evaluating the test results, and building test facilities and instrumentation.

In contrast with one-time initial investment costs, different cost elements are less dominant in periodic recurring costs. Since there is a substantial difference in the resource demands of two kinds of expenditures for the annual operating charges and for the cost associated with the subsystem failures, annual recurring costs are divided into two categories:

(A) The periodic operating costs representing primarily the costs of replacement of subsystem components due to wear and tear, pay and allowances for operation and maintenance crews, fuel and lubricants, and other administrative costs.

(B) The periodic failure costs representing the expenses of repairing or replacing the failed subsystems as well as the loss, if any, suffered by the failure of an entire system owing to the sub-system failure.

Consideration of the kinds of relevant system costs involved in the reliability allocation problem suggests that a U-shaped system cost curve is required. For example, annual operating cost and failure cost are both high when the subsystem failure rate is high, and one-time initial investment is high also when the subsystem failure rate is low. Thus, somewhere between these two extremes, the combined costs are invariably at a minimum.

In order to have a simple relationship between the relevant system costs and allocated failure rates, it is convenient to introduce the variable, relative failure rate, defined as the ratio of base failure rate[1] to allocated failure rate. Relative failure rate for the jth subsystem, x_j, is thus expressed by

$$x_j = \frac{\bar{f}_j}{f_j} \tag{4-1}$$

where $\bar{f}_j$ = base failure rate for the jth subsystem

f_j = allocated failure rate for the jth subsystem

The relative failure rate which gives a measure of departure from the

[1]Base failure rate is obtainable from the existing failure data or manufacturers of the system. MIL-STD-217 provides an extensive list of failure rates for the electronic equipment.

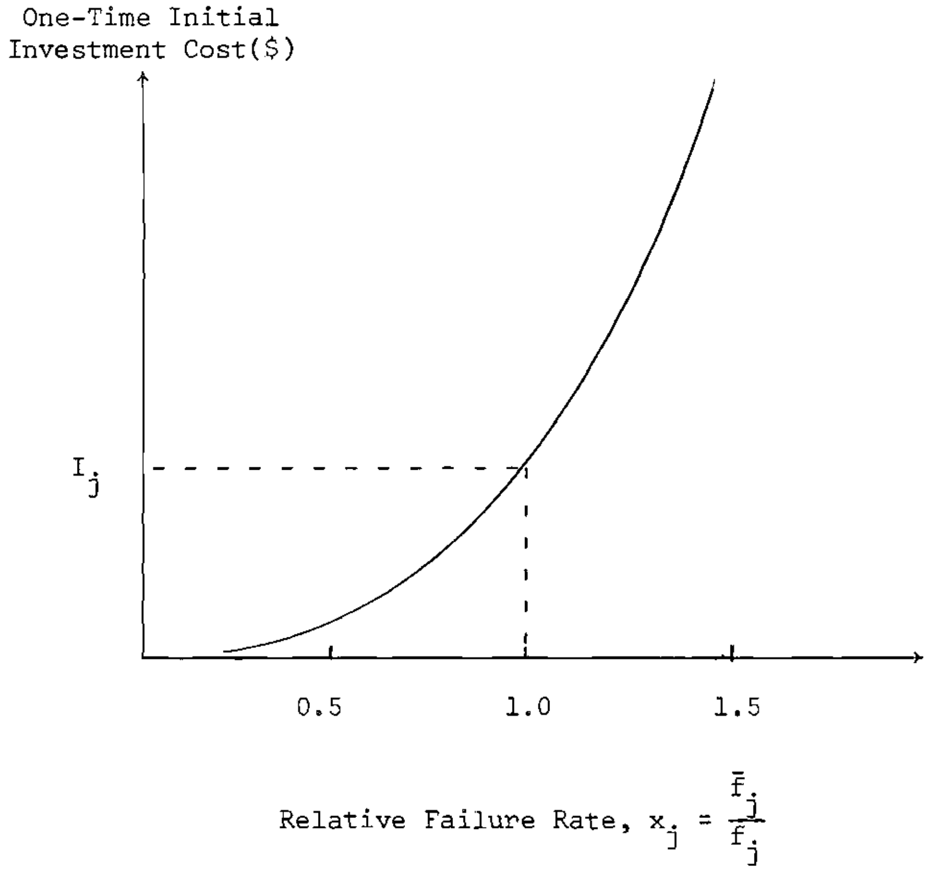


Figure 6. One-Time Initial Investment Cost as a Function of Relative Failure Rate

current state-of-the-art is less than unity when the allocated failure rate is higher than the base failure rate, and is equal to unity when the allocated failure rate is the same as the base failure rate.

When the allocated failure rate for the jth subsystem exceeds the base failure rate whose initial investment costs are known to be I_j dollars, the costs of initial investment tend to decrease below I_j dollars. On the other hand, reduction of subsystem failure rate below its current state-of-the-art may require additional Research and Development activities, more expensive fabrication techniques, better and more expensive components, more extensive quality control and reliability testing; consequently, the initial investment costs tend to increase sharply. It is also conceivable that the amount of initial investment increases more rapidly as the relative failure rate exceeds unity so that the cost curve bends upward.

With these considerations in mind, the relation between the initial investment costs and the relative failure rates may be reasonably indicated by the curve shown in Figure 6, which may be suitably approximated by the following cost function:

$$\text{Initial investment cost for the jth subsystem} = I_j \left(\frac{\bar{f}_j}{f_j}\right)^{S_j} = I_j\, x_j^{S_j} \qquad (4\text{-}2)$$

where x_j represents the relative failure rate of the jth subsystem and the I_j is the initial investment cost when the relative failure rate is unity,[2] and the exponent S_j is a constant which determines the rate at which changes in the cost occur, where $S_j > 1$ for $j = 1,2,\ldots,n$.

Since one-time initial investment costs occur only once throughout the entire service life of a subsystem, it is therefore necessary to multiply the appropriate capital recovery factor to convert this one-time cost into its equivalent annual cost. Equivalent annual cost of the initial investment is thus given by

$$\begin{array}{l}\text{Equivalent annual cost}\\ \text{of initial investment}\end{array} = C_R \, I_j \, x_j^{S_j} \qquad (4\text{-}3)$$

where C_R denotes the capital recovery factor.[3]

In approximation of a cost function, it may be possible to include the fixed cost term in order to improve the approximation in the relevant region. However, it proves to be irrelevant in obtaining optimal allocations since this cost is not dependent upon the allocated failure rate.

Annual operating cost which represents primarily the cost incurred yearly for the operation and routine replacement of the sub-

---

[2]The parameter of the cost function is set at this value since it will give the best possible approximation to the underlying cost curve over the region in which allocated failure rate is expected to fall.

[3]The capital recovery factor is defined as

$$C_R = \frac{i(1+i)^K}{(1+i)^K-1},$$

where K is the expected service life of the subsystem, and i is the interest rate. The value of capital recovery factor may be found in any interest table.

system components due to wear and tear is also affected by the allocated level of subsystem failure rate. Annual operating cost rises rapidly as relative failure rate approaches zero, and gradually decreases at the other extreme where the allocated failure rate is so low that there are infrequent replacements of subsystem components needed.

The relationship between annual operating cost and relative failure rate is shown in Figure 7. This curve may be suitably approximated over a range of interest by a function of the form:

$$\text{Annual operating cost} = \frac{A_j}{\left(\frac{\bar{f}_j}{f_j}\right)} = \frac{A_j}{x_j} \qquad (4\text{-}4)$$

where A_j represents the annual operating cost for the jth subsystem when the relative failure rate is unity.

If the allocated failure rate for the jth subsystem is f_j, over the time T (annual operating time of the system), the jth subsystem can be expected to fail f_jT times. If one lets K_j be the conditional probability that entire system fails given that the jth subsystem has failed,[4] and if the cost associated with a failure of the jth subsystem is Q_j and that involved in a failure of the entire system is Q_S, then the expected cost of subsystem failures over the time period of T hours becomes:

---

[4]A formula for estimating the K_j factor may be given as:

$$K_j = \frac{\text{No. of Entire System Failures Due to jth Subsystem Failures}}{\text{No. of jth Subsystem Failures}}$$

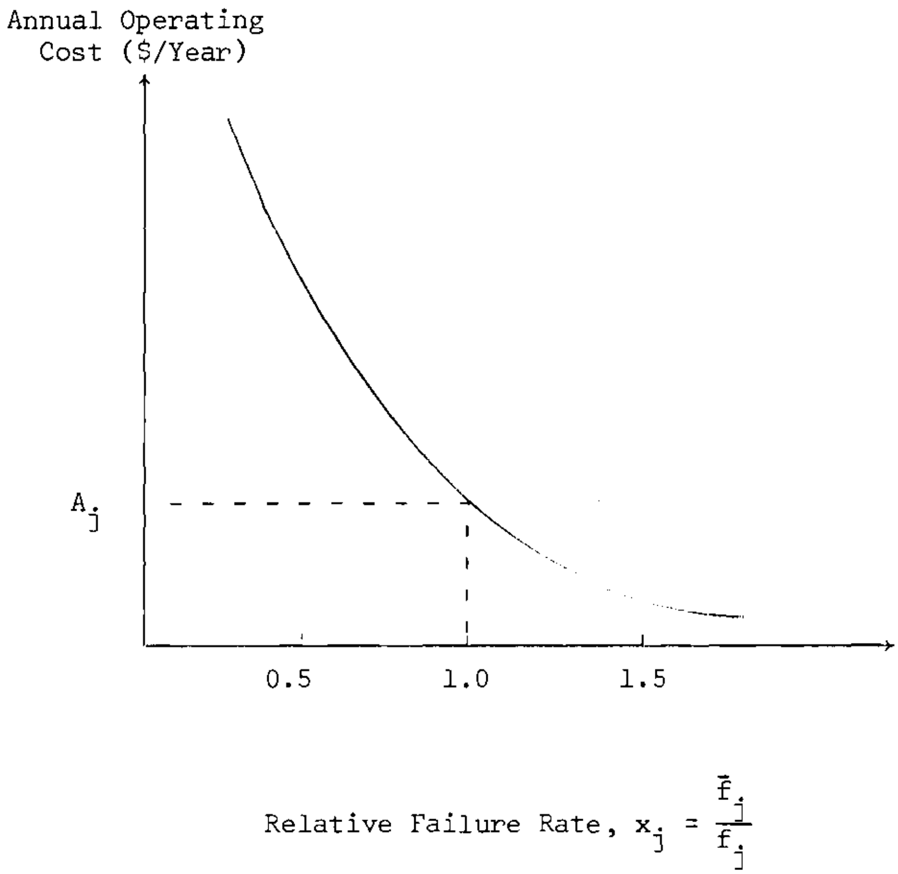


Figure 7. Annual Operating Cost as a Function of Relative Failure Rate

$$\text{Expected cost of subsystem failures} = f_j \, T[K_j \, Q_S + (1 - K_j)Q_j] \tag{4-5}$$

$$= \frac{\bar{f}_j}{x_j} T[Q_j + K_j(Q_S - Q_j)]$$

Derivation of the Allocation Model

Having examined the individual cost elements that depend on the allocated failure rates, it is now possible to construct the total annual cost function for the system reliability allocations. The total annual cost for the jth subsystem is the sum of the relevant cost elements that have been discussed above and is given by:

$$C_j = [C_R \, I_j \, x_j^{\;S_j} \quad \text{(Equivalent Annual Cost of Investments)}$$

$$+ \frac{A_j}{x_j} \qquad \text{(Annual Operating Cost)}$$

$$+ \frac{\bar{f}_j}{x_j} T\{Q_j + K_j(Q_S - Q_j)\}] \quad \text{(Annual Failure Cost)}$$

where, by definition, relative failure rate is given as $x_j = \bar{f}_j/f_j$.

Given the allowable system failure rate corresponding to a specified system reliability requirement, the allocatins of allowable system failure rate among n constituent subsystems are to be accomplished so as to minimize the overall relevant costs. The reliability allocation problem can then be stated formally as:

Find the subsystem failure rates that will minimize the total annual system costs C, where

$$C = \sum_{j=1}^{n} C_j \tag{4-6}$$

and

$$C_j = [C_R\ I_j\ x_j^{S_j} + \frac{A_j}{x_j} + \frac{\bar{f}_j}{x_j} T\{Q_j + K_j(Q_S - Q_j)\}] \tag{4-7}$$

subject to the constraint

$$\sum_{j=1}^{n} \frac{\bar{f}_j}{x_j} = F, \qquad j = 1,2,\ldots,n \tag{4-8}$$

where F is the allowable system failure rate. The optimal subsystem failure rates are those that will minimize the total system costs C. These total costs are shown in Equation (4-6) to be the sum of the n subsystem costs.

The above cost function may be applied to various types of systems simply by using the appropriate numerical values for the cost parameters.[5] It is also to be noted that in the above formulation, no bound on the decision variable, such as nonnegativity constraint of x_j, is placed since for the type of system failure rate allocation

---

[5]The fixed cost terms that do not vary with changes in the allocated failure rate may be added to any of the above cost expressions without affecting the optimal allocations. Costs that are constant are irrelevant in the allocations of system failure rate and hence such cost terms will simply be ignored in the approximation of the cost function.

problem posed here the unconstrained solution will automatically satisfy such a nonnegativity constraint with no exceptions. In minimizing the total system costs, negative failure rates are automatically avoided since these decisions are costly and undesirable. Therefore it is considered adequate not to place this constraint on the solutions.

In order to minimize C subject to the constraint that the allocated subsystem failure rates must add up to the specified system failure rate as shown in (4-8), one forms the Lagrangian function,

$$L = \sum_{j=1}^{n} \left[C_R\, I_j\, x_j^{S_j} + \frac{A_j}{x_j} + \frac{\bar{f}_j}{x_j} T\{Q_j + K_j(Q_S - Q_j)\} \right] + \lambda \left[F - \sum_{j=1}^{n} \frac{\bar{f}_j}{x_j} \right] \qquad (4\text{-}9)$$

where λ is a Lagrange multiplier satisfying the relation of (4-8).

Taking partial derivatives with respect to x_j and λ, and setting them equal to zero, one obtains the first order condition for a minimum:[6]

[6]This first order condition is also sufficient for the absolute minimum since the cost function involved is strictly convex over a region of interest, and therefore any relative minimum of the objective function over the region is also the absolute minimum. The sufficient conditions for a solution to (4-10) and (4-11) to be a local minimum when n = 2 is that

$$\Delta_3 = [-g_{x_1}^2 (f_{x_2x_2} - \lambda g_{x_2x_2}) - g_{x_2}^2 (f_{x_1x_1} - \lambda g_{x_1x_1}) + 2g_{x_1} g_{x_2} (f_{x_1x_2} - \lambda g_{x_1x_2})]$$

must be negative, where f_{x_1} and g_{x_1} denote the partial derivatives of the cost function and constraint function with respect to x_1, respectively, and $\lambda = f_{x_1}/g_{x_1} = f_{x_2}/g_{x_2}$. This condition is satisfied when n = 2 as

$$\Delta_3 = -(\bar{f}_2/x_2^2)^2 C_R I_1 S_1 x_1^{S_1-2}(S_1+1) - \bar{f}_1/x_1^2)^2 C_R I_2 S_2 x_2^{S_2-2}(S_2+1) < 0,$$

and an analogous result will hold for any n.

$$\frac{\partial L}{\partial x_j} = C_R I_j S_j x_j^{S_j - 1} - \frac{A_j}{x_j^2} - \frac{\bar{f}_j}{x_j^2} T[Q_j + K_j(Q_S - Q_j)] + \lambda \frac{\bar{f}_j}{x_j^2} = 0 \qquad (4\text{-}10)$$

$$\frac{\partial L}{\partial \lambda} = F - \sum_{j=1}^{n} \frac{\bar{f}_j}{x_j} = 0 \qquad (4\text{-}11)$$

where $j = 1,2,3,\ldots,n$. The $n + 1$ equations above must be solved simultaneously for the $n + 1$ unknowns, x_j and λ. The relative failure rates so obtained will be functions of the desired system failure rate F. Substituting the solution into the cost function for each subsystem, and summing the costs, one obtains the minimum total annual cost for the specified level of system failure rate.

The solution for the constrained optimal relative failure rate may be written as

$$x_j = \left(\frac{A_j + \bar{f}_j \; T[Q_j + K_j(Q_S - Q_j)] - \lambda \bar{f}_j}{C_R \; I_j \; S_j} \right)^{\frac{1}{S_j + 1}} \qquad (4\text{-}12)$$

where $j = 1,2,\ldots,n$.

In order to use Equation (4-12) to obtain the optimal subsystem failure rate, however, one needs first to solve the system consisting of Equation (4-10) and (4-11) explicitly for the unknown variable λ. The solution would be difficult to compute except in a special case because one desires an expression for λ that does not depend on the allocated subsystem failure rates which are also unknown. Certain methods of solution are presented in the following section.

Computational Methods

Although there will be a unique value of λ which satisfies Equations (4-10) and (4-11), there does not always exist an explicit expression for computing that value.[7] It is therefore desirable to consider a special case which does yield an explicit solution and then to examine the graphical and mathematical approximation which may apply under more general conditions.

Solution for a Special Case

Suppose that the system under consideration will fail in the presence of any subsystem failure, and annual operating costs measured in unit of the base failure rate are the same for all subsystems. That is,

$$K_i = K_j = 1, \text{ and } \frac{A_i}{\bar{f}_i} = \frac{A_j}{\bar{f}_j} = \bar{A}, \tag{4-13}$$

a constant for all i and j.

Under these conditions, Equation (4-12) may be written as:

$$x_j = \left[\frac{\bar{f}_j(\bar{A} + T\, Q_S - \lambda)}{C_R\, I_j\, S_j}\right]^{\frac{1}{S_j+1}} \tag{4-14}$$

---

[7]There always exists a unique solution to the above necessary conditions since, for a given λ, the x_j are uniquely determined, and since the x_j values decrease monotonically as λ is increased, so that there will be only one λ for which (4-11) is satisfied.

where j = 1,2,...,n. Since Equation (4-8) must also hold for the solution, it follows that

$$F = \sum_{j=1}^{n} \frac{\bar{f}_j}{x_j} = \sum_{j=1}^{n} \left(\frac{C_R\ I_j\ S_j\ \bar{f}_j^{\ S_j}}{\bar{A} + T\ Q_S - \lambda} \right)^{\frac{1}{S_j+1}} \tag{4-15}$$

After rearranging the term, one obtains the following expression in λ:

$$\bar{A} + T\ Q_S - \lambda = \frac{1}{F^{S_j+1}} \left[\sum_{j=1}^{n} \sqrt[S_j+1]{C_R\ I_j\ S_j\ \bar{f}_j^{\ S_j}} \right]^{S_j+1} \tag{4-16}$$

Then the x_j are given, from Equation (4-14), by

$$x_j = \frac{\bar{f}_j^{\frac{1}{S_j+1}} \left[\sum_{j=1}^{n} \sqrt[S_j+1]{C_R\ I_j\ S_j\ \bar{f}_1^{\ S_j}} \right]}{F \left[C_R\ I_j\ S_j \right]^{\frac{1}{S_j+1}}} \tag{4-17}$$

where j = 1,2,...,n. Hence, the allocated failure rate for the jth subsystem is given by

$$f_j = \frac{\bar{f}_j}{x_j} = \frac{F \cdot \sqrt[S_j+1]{C_R\ I_j\ S_j\ \bar{f}_j^{\ S_j}}}{\sum_{j=1}^{n} \sqrt[S_j+1]{C_R\ I_j\ S_j\ \bar{f}_j^{\ S_j}}} \tag{4-18}$$

where j = 1,2,...,n.

Since the above expression no longer depends on the unknown λ, the optimal failure rate for each subsystem will now be a constant fraction of the system failure rate.

From Equation (4-12), the unconstrained relative failure rate for the jth subsystem, x_j^*, is obtained by setting λ equal to zero as:

$$x_j^* = \left(\frac{A_j + \bar{f}_j\ T\ [Q_j + K_j(Q_S - Q_j)]}{C_R\ I_j\ S_j}\right)^{\frac{1}{S_j+1}} \tag{4-19}$$

Then unconstrained failure rate for the jth subsystem is

$$f_j^* = \frac{\bar{f}_j}{x_j^*} = \left(\frac{C_R\ I_j\ S_j\ \bar{f}_j^{\ S_j+1}}{A_j + \bar{f}_j\ T[Q_j + K_j(Q_S - Q_j)]}\right)^{\frac{1}{S_j+1}} \tag{4-20}$$

Forming the ratio of $\frac{f_j}{f_j^*}$, that is, the ratio between the constrained failure rate and the unconstrained failure rate for the jth subsystem, one obtains:

$$\frac{f_j}{f_j^*} = \frac{\dfrac{F\left(C_R\ I_j\ S_j\ \bar{f}_j^{\ S_j}\right)^{\frac{1}{S_j+1}}}{\sum\left(C_R\ I_j\ S_j\ \bar{f}_j^{\ S_j}\right)^{\frac{1}{S_j+1}}}}{\left(\dfrac{C_R\ I_j\ S_j\ \bar{f}_j^{\ S_j+1}}{A_j + \bar{f}_j\ T[Q_j + K_j(Q_S-Q_j)]}\right)^{\frac{1}{S_j+1}}} \tag{4-21}$$

$$= \frac{F\left[A_j + \bar{f}_j\, T[Q_j + K_j(Q_S - Q_j)]\right]^{\frac{1}{S_j+1}}}{\sum\left(C_R\, I_j\, S_j\, \bar{f}_j^{\,S_j+1}\right)^{\frac{1}{S_j+1}}}$$

But, from Equation (4-20), it follows that

$$\left(C_R\, I_j\, S_j\, \bar{f}_j^{\,S_j+1}\right)^{\frac{1}{S_j+1}} = f_j^*[A_j + \bar{f}_j\, T\{Q_j + K_j(Q_S - Q_j)\}]^{\frac{1}{S_j+1}}$$

Substituting the above equation into (4-21), one obtains the following simple relation:

$$\frac{f_j}{f_j^*} = \frac{F\left(A_j + \bar{f}_j\, T[Q_j + K_j(Q_S-Q_j)]\right)^{\frac{1}{S_j+1}}}{\sum\limits_j f_j^*\left(A_j + \bar{f}_j\, T[Q_j + K_j(Q_S-Q_j)]\right)^{\frac{1}{S_j+1}}} = \frac{F}{\sum\limits_j f_j^*} \qquad (4\text{-}23)$$

or

$$f_j = \frac{F}{\sum\limits_j f_j^*} \cdot f_j^* \qquad (4\text{-}24)$$

where $j = 1,2,\ldots,n$. The above equation shows that each constrained subsystem failure rate f_j is directly proportional to the value of the unconstrained failure rate f_j^* with the proportionality constant equal to $F/\sum\limits_j f_j^*$.

Equation (4-24) is valid only when the assumption given in (4-13) holds. Except in a special case such as the above it is not always possible to obtain a simple solution. One can, however, use the following graphical method for determining the correct value of λ which does satisfy the given constraint of Equation (4-8).

1. First select some numerical value of λ. One guide in selecting reasonable values for λ is to notice that it is the marginal cost of system failure rate.[8]

2. Compute x_j, where $j = 1,2,\ldots,n$, from Equation (4-12) for each value of λ.

3. Compute $F(\lambda) = \sum_j \bar{f}_j/x_j(\lambda)$ for each value of λ (Equation (4-8)).

4. Plot the points so obtained with coordinates λ, $F(\lambda)$, and draw a smooth curve through the plotted points, as illustrated in Figure 8.

5. Finally interpolate graphically to obtain the desired value of λ satisfying the specified value of F.

The construction of the graph may be facilitated by first locating the asymptotes. From Equations (4-12) and (4-8), it is clear that as F approaches zero, the quantity $A_j + \bar{f}_j\ T\{Q_j + K_j(Q_S-Q_j)\} - \lambda\bar{f}_j$ approaches infinity; and hence, λ approaches minus infinity. Similarly, for very large values of F, at least one of the above quantities will approach zero, and none of them will be negative. Therefore, λ will approach the minimum of $\frac{1}{\bar{f}_j}\{A_j + \bar{f}_j\ T[Q_j+K_j(Q_S-Q_j)]\}$

[8]This statement is verified in the following section.

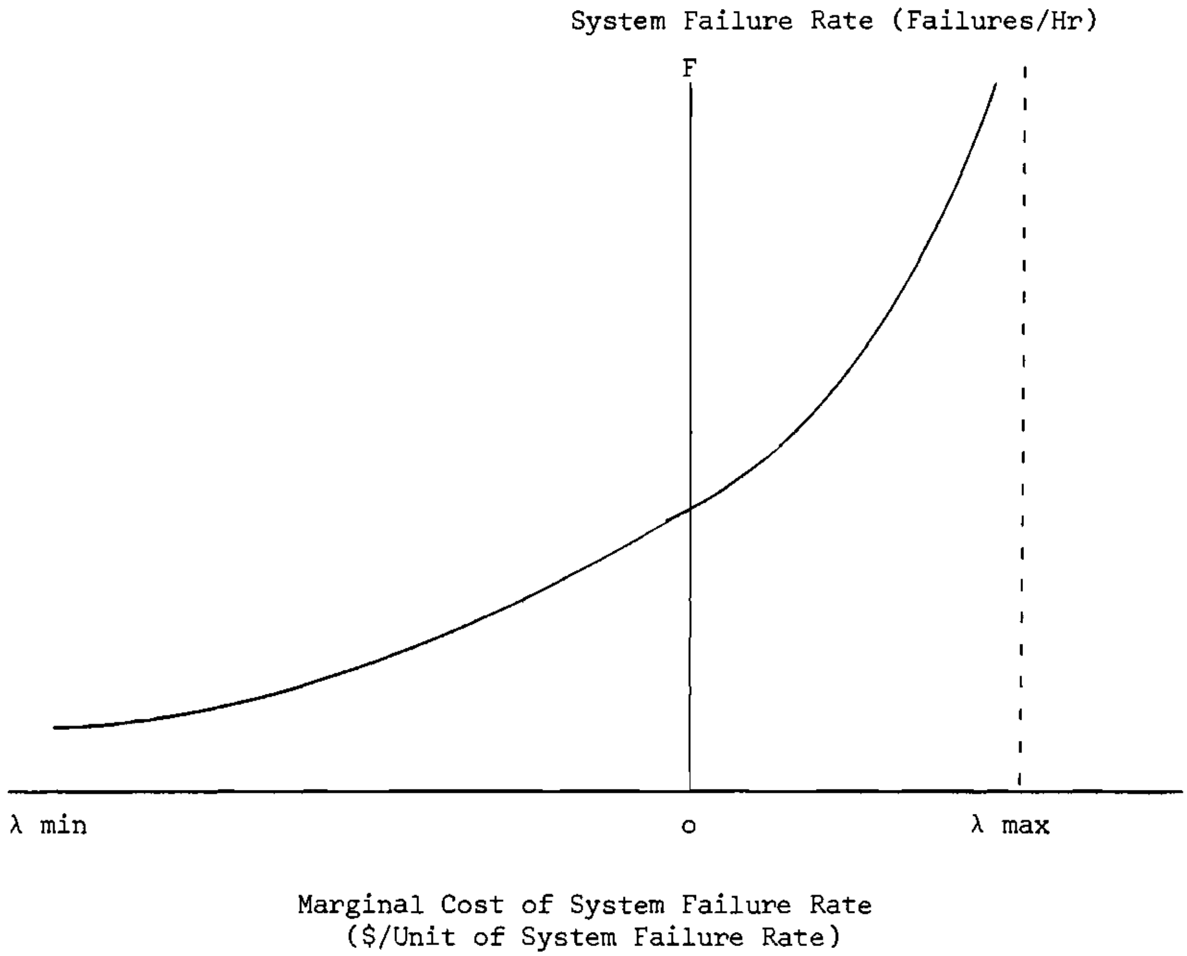


Figure 8. System Failure Rate as a Function of Marginal Cost of System Failure Rate, λ.

for $j = 1,2,\ldots,n$. This accounts for two asymptotes shown in Figure 8.

Once this curve has been constructed, it may be used repeatedly as long as the cost parameters are unchanged. In order to determine the optimum allocated failure rate for each subsystem, one needs only to obtain from the graph the value of λ corresponding to the specified level of system failure rate, and insert this value in Equation (4-12) to determine x_j for each subsystem. The allocated failure rate for the jth subsystem is then obtained from the relation $f_j = \bar{f}_j/x_j$. Substituting these values of f_j into Equation (4-6) gives the minimum total annual system costs when the allowable system failure rate is optimally allocated among constituent subsystems.

Linear Approximation in λ

The condition for minimum cost given in Equation (4-10) may be approximated in the desired region by a linear function of the system failure rate. The linear approximation is found from the differential changes in the Lagrange multiplier, $d\lambda$, corresponding to the differential changes in the system failure rate, dF.

Since the minimum cost condition is to be satisfied, the total differentials dx_j and $d\lambda$ must, from Equation (4-10), satisfy:

$$\left[C_R\ I_j\ S_j(S_j-1)\ x_j^{S_j-2} + \frac{2A_j}{x_j^3} + \frac{2\bar{f}_j}{x_j^3} T\{Q_j + K_j(Q_S-Q_j)\} - \frac{2\bar{f}_j\lambda}{x_j^3}\right] dx_j + \frac{\bar{f}_j}{x_j^2} d\lambda = 0 \qquad (4\text{-}25)$$

for $j = 1,2,\ldots,n$. Solving for dx_j gives

$$dx_j = - \frac{x_j \bar{f}_j}{C_R I_j S_j (S_j - 1) x_j^{S_j+1} + 2A_j + 2\bar{f}_j T[Q_j + K_j(Q_S - Q_j)] - 2\lambda \bar{f}_j} d\lambda \quad (4\text{-}26)$$

Since the system failure rate constraint of Equation (4-8) must also be satisfied, substituting the above equation into the constraint equation

$$dF = - \sum_{j=1}^{n} \frac{\bar{f}_j}{x_j^2} dx_j \quad (4\text{-}27)$$

to obtain

$$dF = \sum_{j=1}^{n} \left(\frac{\bar{f}_j}{x_j}\right) \cdot \frac{\bar{f}_j \, d\lambda}{C_R I_j S_j (S_j - 1) x_j^{S_j+1} + 2A_j + 2\bar{f}_j T[Q_j + K_j(Q_S - Q_j)] - 2\lambda \bar{f}_j} \quad (4\text{-}28)$$

But, from Equation (4-10), one obtains the following relation:

$$\lambda = - \frac{1}{\bar{f}_j} \cdot C_R I_j S_j x_j^{S_j+1} + \frac{A_j}{\bar{f}_j} + T[Q_j + K_j(Q_S - Q_j)] \quad (4\text{-}29)$$

Substituting the above equation into (4-28), and simplifying the terms, one obtains:

$$dF = \sum_{j=1}^{n} \frac{\bar{f}_j^2}{C_R I_j S_j (S_j + 1) x_j^{S_j+2}} d\lambda \quad (4\text{-}30)$$

Solving for the changes in the multiplier then gives:

$$d\lambda = \left[\sum_{j=1}^{n} \frac{\bar{f}_j^2}{C_R\ I_j\ S_j(S_j+1)\ x_j^{S_j+2}} \right]^{-1} dF \tag{4-31}$$

The above equation makes it possible to determine the value of the Lagrange multiplier corresponding to a value of the specified system failure rate. The linear terms of the Taylor's series expansion are obtained by replacing the differentials by finite differences from the point of expansion. Thus replacing the differentials in Equation (4-31) with finite differences from the point of expansion, one finds the linear expression for λ:

$$\lambda = \lambda^o + \left[\sum_{j=1}^{n} \frac{\bar{f}_j^2}{C_R\ I_j\ S_j(S_j+1)\ (x_j^o)^{S_j+2}} \right]^{-1} (F - F^o) \tag{4-32}$$

where the superscript ^{o} denotes the value of the variable evaluated at the point of expansion.

Once the value of λ^o is selected, the corresponding values of x_j^o and F^o in Equation (4-32) can be found from (4-12) and (4-8), respectively. The first convenient point of expansion may be $\lambda^o = 0$, that is, the unconstrained point of x_j^o and F^o, given by Equations (4-12) and (4-8), respectively. This approximation amounts to using the line tangent to the curve at the point of expansion λ^o.

The above linear approximation method is analogous to Newton's method of finding the solution to the equation of the form $f(x) = 0$.

Newton's method of solution to (4-8) and (4-12) is an iterative procedure based on the replacement of a curve by a straight line, the tangent, in the neighborhood of an initial estimate of λ, which will satisfy the constraint in (4-8). Suppose that in a given situation, the unconstrained system failure rate exceeds the allowable system failure rate F^* as shown in Figure 9, and furthermore assume that λ^o is an initial estimate of λ which will satisfy the constraint in (4-8).

Then, Newton's method calls for replacing the marginal cost curve by the tangent F^oA which intersects the $F(\lambda) = F^*$ line at the point A. From the definition of the derivative at λ^o,

$$F'(\lambda^o) = \frac{F^o - F^*}{\lambda^o - \lambda_1} \tag{4-33}$$

But, from Equation (4-30), one has

$$F'(\lambda^o) = \sum_{j=1}^{n} \frac{\bar{f}_j^{\,2}}{C_R\, I_j\, S_j(S_j+1)\, (x_j^o)^{S_j+2}}$$

and therefore,

$$\lambda_1 = \lambda^o + \left[\sum_{j=1}^{n} \frac{\bar{f}_j^{\,2}}{C_R\, I_j\, S_j(S_j+1)\, (x_j^o)^{S_j+2}}\right]^{-1} (F^* - F^o) \tag{4-34}$$

The same procedure can be applied at the point $F_1(\lambda_1)$ and thus leads to

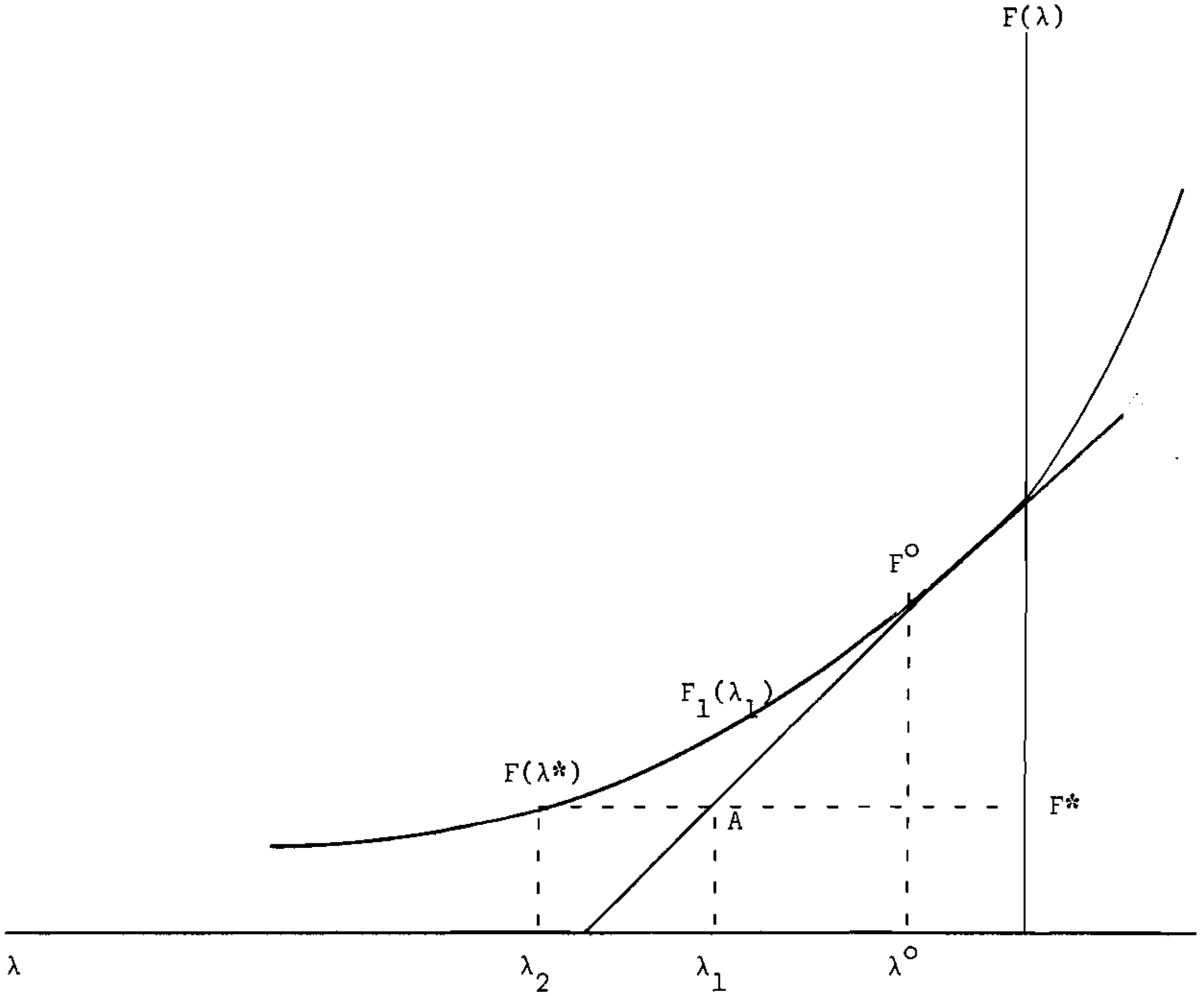


Figure 9. Marginal Cost of System Failure Rate

a better approximation λ_2:

$$\lambda_2 = \lambda_1 + \left(\sum_{j=1}^{n} \frac{\bar{f}_j^2}{C_R\ I_j\ S_j(S_j+1)\ x_j(\lambda_1)^{S_j+2}} \right)^{-1} (F^* - F_1(\lambda_1)) \qquad (4\text{-}35)$$

This iterative process will continue until the desired solution is finally found. The above Equation (4-35) is identical to the linear approximation found in (4-32).

It is to be noted that the estimate of λ obtained by using the linear approximation Equation (4-32) will always exceed the true value of λ which corresponds to a specified value of F. This is simply because the line tangent to the marginal cost curve at the point of expansion always lies to the right of the curve so that linear approximation Equation (4-32) overstates that value of λ corresponding to a specified F (see Figure 9).

It is also seen from the figure that the unconstrained point where $\lambda^o = 0$ is not always an appropriate point of expansion. As the value of the system failure rate constraint F^* differs greatly from the unconstrained point where $\lambda^o = 0$, the estimate of λ^* obtained from (4-32), using $\lambda^o = 0$ as the point of expansion gives a poor approximation to the exact value of λ^*. However, successive applications of (4-32) in this case will converge toward the exact value of λ^* rapidly as the shape of the curve suggests.

Given the desired value of F, Equation (4-32) would lead to a linear approximation in λ, the marginal cost of system failure rate when the system failure rate is allocated optimally among all the subsystems. Substituting this value of λ into Equation (4-12), the optimal subsystem failure rate is readily found for each subsystem.

It is sometimes convenient to use linear approximation for x_j as well. In this case, the solution to the unconstrained problem (that is, $\lambda = 0$) provides a convenient point from which linear approximation of the functions in Equation (4-12) may be taken. The linear approximation to Equation (4-12) can be found by replacing the differentials in Equation (4-26) with finite differences:

$$x_j = x_j^o - \frac{x_j^o \, \bar{f}_j(\lambda - \lambda^o)}{C_R \, I_j \, S_j(S_j - 1)(x_j^o)^{S_j+1} + 2A_j + 2\bar{f}_j \, T[Q_j + K_j(Q_S - Q_j)] - 2\bar{f}_j \, \lambda^o} \tag{4-36}$$

From Equation (4-12) one has:

$$(x_j^o)^{S_j+1} = \frac{A_j + \bar{f}_j \, T[Q_j + K_j(Q_S - Q_j)] - \lambda^o \, \bar{f}_j}{C_R \, I_j \, S_j} \tag{4-37}$$

Substituting the above equation into (4-36) and simplifying, one obtains:

$$x_j = x_j^o - \frac{x_j^o \, \bar{f}_j \, (\lambda - \lambda^o)}{(S_j + 1) \, [A_j + \bar{f}_j \, T\{Q_j + K_j(Q_S - Q_j)\}] - \lambda^o \, \bar{f}_j} \tag{4-38}$$

where $j = 1,2,\ldots,n$. Now the value of λ can be calculated, given the desired F. Substituting this value of λ into Equation (4-38), optimal x_j are obtained for each subsystem.

Interpretation of the Lagrange Multiplier

It has been stated that the Lagrange multiplier introduced is the derivative of the minimum cost function with respect to F. In general, the optimum value of x_j and λ will depend on the particular value of F which appears in a given allocation problem. Therefore the minimum value of C in the cost Equation (4-6) may be regarded as a function of the parameter, F. Then the differential changes in the system failure rate, dF, would change total annual cost shown in Equation (4-6) by dC because of the resulting differential changes dx_j:

$$dC = \sum_{j=1}^{n} \left[C_R\, I_j\, S_j\, x_j^{S_j - 1} - \frac{A_j}{x_j^2} - \frac{\bar{f}_j}{x_j^2} T\{Q_j + K_j(Q_S - Q_j)\} \right] dx_j \qquad (4\text{-}39)$$

Since the first order condition for minimum given in Equation (4-10) must always be satisfied at the optimum solution, it follows from Equation (4-10) that

$$\left[C_R\, I_j\, S_j\, x_j^{S_j - 1} - \frac{A_j}{x_j^2} - \frac{\bar{f}_j}{x_j^2} T\{Q_j + K_j(Q_S - Q_j)\} \right] dx_j = -\lambda \frac{\bar{f}_j}{x_j^2} dx_j \qquad (4\text{-}40)$$

where $j = 1,2,\ldots,n$.

Substituting the above relation into (4-39) gives

$$dC = \sum_{j=1}^{n} -\lambda \cdot \frac{\bar{f}_j}{x_j^2} dx_j \qquad (4\text{-}41)$$

But from Equation (4-8), it follows that

$$\lambda \sum_{j=1}^{n} (-1) \cdot \frac{\bar{f}_j}{x_j^2} dx_j = \lambda \cdot dF$$

so that

$$dC = \frac{\partial C^*}{\partial F} dF = \lambda \cdot dF \tag{4-42}$$

The above observation shows that the partial derivative of the minimum cost function with respect to F is simply equal to the Lagrange multiplier λ, evaluated at F. In this case, the physical dimension of λ is dollars per unit of system failure rate. Therefore λ may be regarded as the marginal cost of system failure rate optimally allocated among subsystems. Generally speaking, this value of λ itself indicates how much the minimum cost will be changed if the allowable system failure rate is increased by one unit.

A Numerical Example

The following hypothetical example will illustrate how the allocation method developed may be applied to assign a numerical failure rate to each of constituent units in a given system. The system under consideration consists of ten subsystems whose expected service life is assumed to be ten years. For the purpose of comparison, the capital recovery factor C_R is taken to be 0.1295 (5 per cent--ten years), and the total annual operating hours are 8760 hours. The cost associated with a failure of the entire system is assumed to be \$1000, and the system failure rate requirement is specified to be 0.024 failures per hour. The other relevant input data are summarized in Table 8.

The allocation problem is then to determine the subsystem failure rates that will minimize the annual system costs. To begin the analysis one will first find the solutions to Equations (4-10), for $j = 1,2,\ldots,10$, ignoring the specified constraint of (4-8). Using Equation (4-12) by letting $\lambda = 0$, and the relation $f_j = \bar{f}_j/x_j$, one will find the following subsystem failure rates for the case of the unconstrained problem:

$$f_1 = 3007 \times 10^{-6} \qquad f_2 = 2002 \times 10^{-6}$$
$$f_3 = 1601 \times 10^{-6} \qquad f_4 = 4592 \times 10^{-6}$$
$$f_5 = 2004 \times 10^{-6} \qquad f_6 = 1201 \times 10^{-6}$$
$$f_7 = 2503 \times 10^{-6} \qquad f_8 = 1001 \times 10^{-6}$$
$$f_9 = 3501 \times 10^{-6} \qquad f_{10} = 3998 \times 10^{-6}$$

and

$$\sum_{j=1}^{10} f_j = 25410 \times 10^{-6} ,$$

which would exceed the specified system failure rate requirement. Thus the constraint will be active in this case. It is therefore necessary to find the solutions explicitly to Equations given in (4-10) and $\sum_{j=1}^{10} \bar{f}_j/x_j = 24000 \times 10^{-6}$.

As suggested in the previous section, a convenient method of solution in this case is to select a $\lambda < 0$, determining the unique x_j

Table 8. Numerical Input Data for Example

Subsystem Number J	$\bar{f}_j$ (Failures/Hr)	I_j (\$)	S_j	A_j (\$)	E_j	Q_j (\$/Failure)
1	0.0030	195,000	2	26,000	0.9	200
2	0.0020	300,000	2	60,000	1.0	400
3	0.0016	350,000	2	77,000	0.9	550
4	0.0046	170,000	2	10,000	0.8	250
5	0.0020	200,000	2	35,000	0.9	400
6	0.0012	250,000	2	54,000	1.0	600
7	0.0025	250,000	2	44,000	0.9	350
8	0.0010	118,000	2	21,700	1.0	250
9	0.0035	150,000	2	10,000	0.9	400
10	0.0040	270,000	2	35,000	1.0	300

C_R = 0.1295 (5% - 10 Years)

T = 8760 (Hours/Year)

Q_S = 1000 (\$/Failure)

values from (4-12), and compute the quantity $\sum_{j=1}^{n} \bar{f}_j/x_j$. If $\sum_{j=1}^{n} \bar{f}_j/x_j > 24000 \times 10^{-6}$, one then selects a smaller value of λ and repeat the computations. If $\sum \bar{f}_j/x_j < 2400 \times 10^{-6}$, one selects a larger value of λ and repeat the computations. In any case there will be a unique solution to (4-10) and (4-11) since, for a given λ, the x_j are uniquely determined, and since the values of x_j decrease monotonically as λ is increased, there will exist only one λ for which $\sum \bar{f}_j/x_j = 24000 \times 10^{-6}$.

As described in the previous section, one proceeds to compute the x_j values for the selected values of λ and plot the results of the computation to obtain a curve showing the overall system failure rate as a function of Lagrange multiplier. The results of these iterative computations beginning with $\lambda = -1.0 \times 10^6$ are shown in Table 9. These results are plotted in Figure 10, showing a relationship between $F(\lambda)$ and λ, from which one may also obtain a graphical solution to λ. In this case, one can easily obtain the appropriate value of λ, that is, $\lambda = -3.2 \times 10^6$.

Then the optimum allocated failure rates and minimum total annual system cost are found as:

$$
\begin{array}{ll}
f_1 = 2835 \times 10^{-6} & f_2 = 1948 \times 10^{-6} \\
f_3 = 1572 \times 10^{-6} & f_4 = 4172 \times 10^{-6} \\
f_5 = 1926 \times 10^{-6} & f_6 = 1178 \times 10^{-6} \\
f_7 = 2406 \times 10^{-6} & f_8 = 967 \times 10^{-6} \\
f_9 = 3216 \times 10^{-6} & f_{10} = 3780 \times 10^{-6}
\end{array}
$$

Table 9. Computational Results for Example

λ	Allocated Subsystem Failure Rate, f_j (in Units of 10^{-6}) — Subsystem Number 1	2	3	4	5	6	7	8	9	10	$F = \sum_{j=1}^{10} f_j$
-1.0×10^6	2949	1984	1592	4443	1979	1194	2472	990	3401	3924	24931×10^{-6}
-2.0×10^6	2895	1968	1583	4312	1955	1187	2442	980	3313	3856	24492×10^{-6}
-2.5×10^6	2870	1960	1579	4252	1943	1183	2428	975	3272	3824	24286×10^{-6}
-3.0×10^6	2845	1951	1574	4194	1931	1179	2412	970	3231	3791	24078×10^{-6}
-3.2×10^6	2835	1948	1572	4172	1926	1177	2407	967	3216	3780	24000×10^{-6}
-3.8×10^6	2808	1940	1567	4109	1914	1174	2391	962	3173	3744	23786×10^{-6}
-4.0×10^6	2799	1937	1565	4089	1910	1173	2386	960	3159	3732	23713×10^{-6}
-4.6×10^6	2772	1928	1560	4031	1897	1169	2370	955	3118	3698	23500×10^{-6}
-5.0×10^6	2755	1922	1556	3994	1889	1166	2360	952	3092	3677	23365×10^{-6}
-6.0×10^6	2714	1907	1548	3907	1869	1159	2334	942	3030	3623	23039×10^{-6}

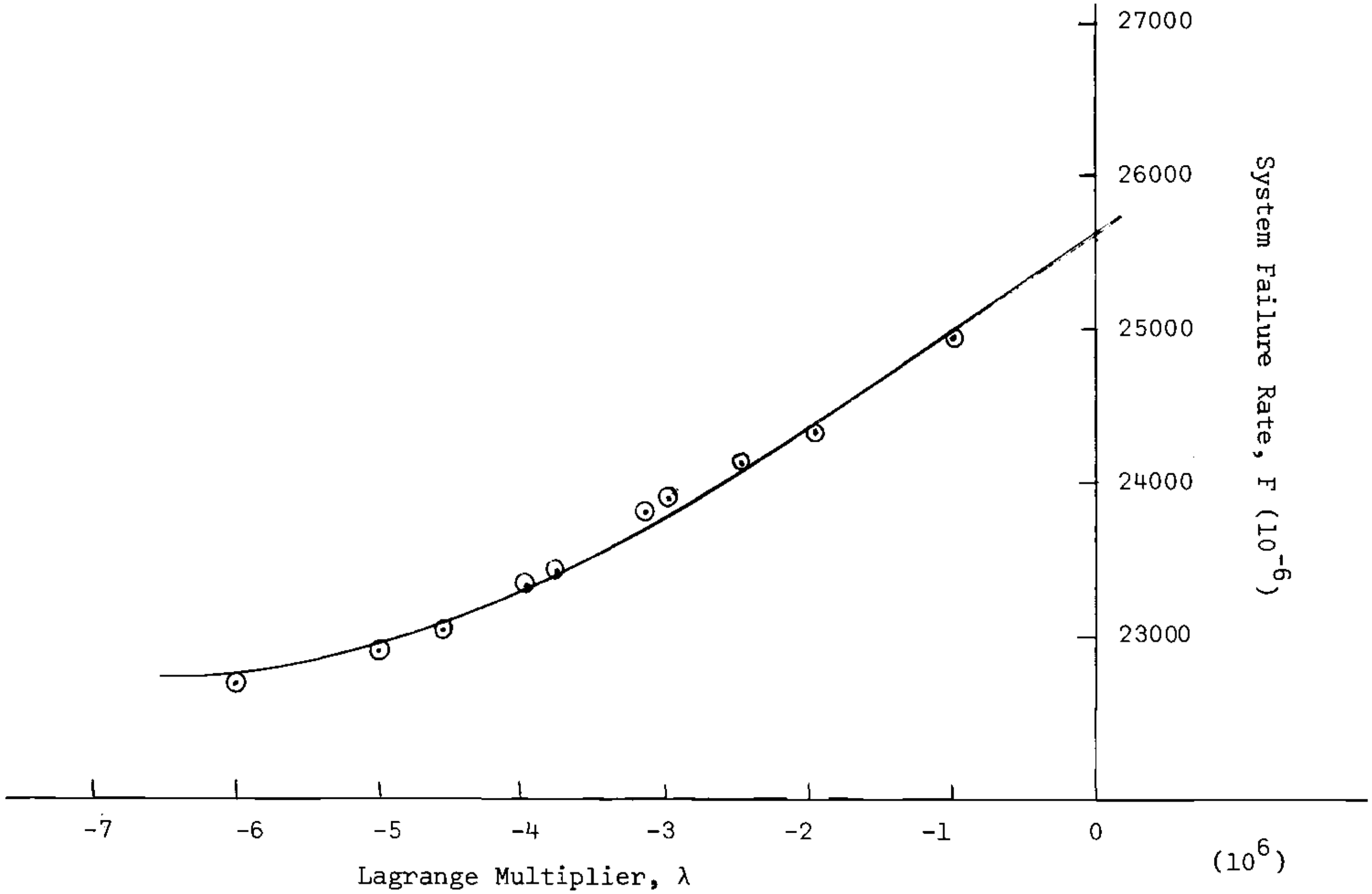


Figure 10. System Failure Rate as a Function of Lagrange Multiplier

and

$$\text{System Failure Rate} = \sum_{j=1}^{10} f_j = 24000 \times 10^{-6}$$

$$\text{Minimum Total System Cost} = 876{,}012$$

Instead of searching by trial and error method a correct value of λ which will satisfy the given constraint, one may equally use the linear approximation method as described in the previous section. The linear approximation equation for λ is given in (4-32). In order to apply the linear approximation method, one must first select the appropriate parameters appearing in (4-32). Table 10 summarizes the calculations of required parameters in Equation (4-32). In evaluating these parameters in (4-32), the unconstrained subsystem failure rate for which $\lambda^o = 0$ is taken as the point of expansion. The corresponding x_j^o is then obtained from (4-12) by letting $\lambda = 0$, and the unconstrained subsystem failure rates are given by the relation of (4-1). The unconstrained system failure rate is then the total of f_j^o so obtained in column 4: $F^o = 25410 \times 10^{-6}$. The last column of the Table 10 then gives:

$$\sum_{j=1}^{10} \frac{\bar{f}_j^2}{C_R\ I_j\ S_j(S_j+1)\ (x_j^o)^{S_j+2}} = 470 \times 10^{-12}$$

Substituting into Equation (4-32), one then finds the linear approximation for λ as:

Table 10. Calculation of the Parameters in the Linear Approximation Formula

Sub-system J	$\bar{f}_j$	x_j^o	f_j^o	$\bar{f}_j^2/C_R I_j S_j(S_j+1)(x_j^o)^{S_j+2}$
1	3000×10^{-6}	0.9977	3007×10^{-6}	54×10^{-12}
2	2000×10^{-6}	0.9991	2002×10^{-6}	15×10^{-12}
3	1600×10^{-6}	0.9993	1601×10^{-6}	8×10^{-12}
4	4600×10^{-6}	1.0018	4592×10^{-6}	150×10^{-12}
5	2000×10^{-6}	0.9980	2004×10^{-6}	23×10^{-12}
6	1200×10^{-6}	0.9999	1201×10^{-6}	7×10^{-12}
7	2500×10^{-6}	0.9988	2503×10^{-6}	31×10^{-12}
8	1000×10^{-6}	0.9998	1001×10^{-6}	10×10^{-12}
9	3500×10^{-6}	0.9999	3501×10^{-6}	101×10^{-12}
10	4000×10^{-6}	1.0005	3998×10^{-6}	71×10^{-12}
Total			25410×10^{-6}	470×10^{-12}

$$\lambda = \lambda^{o} + \left[\sum_{j=1}^{10} \frac{\bar{f}_j^2}{C_R \; I_j \; S_j(S_j+1) \; (x_j^{o})^{S_j+2}} \right]^{-1} (F - F^{o})$$

$$= \frac{10^{12}}{470} (24000 \times 10^{-6} - 25410 \times 10^{-6})$$

$$= - 3.0 \times 10^{6}$$

It is to be noted from Table 9 that the above estimate of λ exceeds the correct value of $\lambda(-3.2 \times 10^{6})$ which corresponds to the specified $F = 24000 \times 10^{-6}$. This is because of the fact that the line tangent to the curve at the point of expansion always lies to the right of the curve, so that Equation (4-32) overstates the value of λ corresponding to $F = 24000 \times 10^{-6}$. However, if this linear approximation for λ is applied once more, making $\lambda^{o} = -3.0 \times 10^{6}$ as the new point of expansion one may readily obtain the desired value of λ.

From Table 9, the allocated subsystem failure rates for $\lambda = -3.0 \times 10^{6}$ are found as:

$f_1 = 2845 \times 10^{-6}$	$f_2 = 1951 \times 10^{-6}$
$f_3 = 1574 \times 10^{-6}$	$f_4 = 4194 \times 10^{-6}$
$f_5 = 1931 \times 10^{-6}$	$f_6 = 1179 \times 10^{-6}$
$f_7 = 2412 \times 10^{-6}$	$f_8 = 970 \times 10^{-6}$
$f_9 = 3231 \times 10^{-6}$	$f_{10} = 3791 \times 10^{-6}$

and

$$\sum_{j=1}^{10} f_j = 24078 \times 10^{-6}$$

The values of the parameters in (4-32) at the point of expansion $\lambda^{o} = -\ 3.0 \times 10^{6}$ are:

J	$\bar{f}_j^2/C_R\ I_j\ S_j(S_j+1)(x_j^o)^{S_j+2}$
1	48×10^{-12}
2	16×10^{-12}
3	9×10^{-12}
4	111×10^{-12}
5	22×10^{-12}
6	8×10^{-12}
7	28×10^{-12}
8	10×10^{-12}
9	76×10^{-12}
10	62×10^{-12}
Total =	390×10^{-12}

Substituting into (4-32), one then finds the second linear estimate of λ as:

$$\lambda = \lambda^{o} + \left(\sum_{j=1}^{10} \frac{\bar{f}_j^2}{C_R\ I_j\ S_j(S_j+1)\ (x_j^o)^{S_j+2}} \right)^{-1} (F - F^{o})$$

$$= -\ 3.0 \times 10^{6} + \frac{10^{12}}{390}\,(2400 \times 10^{-6} - 24078 \times 10^{-6})$$

$$= -\ 3.2 \times 10^{6}$$

The second approximation of λ improves the approximation far better and yields the desired value of λ satisfying the given constraint.

Concluding Remarks

In this chapter, a reliability allocation method based on the total system costs has been developed. Since every new situation arisen in the reliability allocation process will call for modification of some of the cost estimating relationships, the methodology presented here is intended to be sufficiently general to apply to a variety of situations.

One of the crucial tasks in applying this allocation method is the determination from historical data of useful relationships between the relevant system costs and the levels of allocated failure rate. Because there are insufficient data for estimating separately the many cost elements involved, they are grouped as an aggregate into three major cost elements--one-time initial investments, annual operating cost, and annual failure cost of the subsystem.

Of all cost elements involved in the reliability allocation process, the cost of initial investments is probably subject to most errors since this cost involves a great deal of uncertainty in the estimations. This uncertainty about cost estimates is merely a reflection of the uncertainties inherent in the Research and Development, prototype fabrication, and testing processes for a newly-developed system. However, this limitation is not considered a permanent one since more and more data will soon become available from systems now under development and from those developed in the future.

The allocation method presented here has several characteristics. First, it takes into account many of the prime cost elements that are relevant in the reliability allocatin, and attempts to minimize the overall system costs in the allocation process. Many of these relevant cost elements are neglected or omitted in existing methods used to allocate system reliability requirement. Secondly, it makes a clear distinction between the one-time initial investment cost and recurring operating expenses which will not only permit an easy measurement of cost elements but also actually minimize the errors associated with the cost estimations. Thirdly, the computational scheme used in the allocation method is rather simple so that the solutions are easy to obtain in a straightforward manner. The method also provides a basis for evaluating alternative system designs in terms of both cost of a given level of system reliability and system reliability obtainable for a fixed cost. The method is "mixed" in a sense that it involves allocations of system reliability requirement as well as planning decisions.

CHAPTER V

RELIABILITY IMPROVEMENT THROUGH REPAIRABLE REDUNDANCY

Preliminary Considerations

In any system design, system reliability requirements cannot be met simply by designing the system that will perform the intended functions. When a preliminary design concept has been proposed, it is necessary to determine how consistent it is with the desired reliability and other performance requirements before time and funds are spent in the detailed design and development. Comparison of pre-design predictions with the specified system requirements is always necessary at this early stage of the development program because it provides system designers with the magnitude and causes of discrepancies between predictions and requirements, and helps to select a proper design approach to satisfy the requirements.

If a reliability requirement greatly exceeds the predicted value, it may be logical to decide that this design concept is not practically feasible. In this situation the system designer is faced with the problem of developing a new design concept which may require an entirely different functional approach to meet the requirements. The usual approaches employed by designers for improving system reliability include such methods as redundancy applications at critical elements, design simplification and derating techniques, use of improved components, and better production and quality control.

Of these approaches toward the reliability design, redundancy application is considered by many system designers as one of the easiest and most appropriate means of obtaining reliability improvement. For a certain redundant system in which system elements are assumed to be independent along with perfect reliability of standby elements and failure-free switching devices, redundancy applications at weak system elements will certainly improve the system reliability significantly. However, in many cases of redundancy applications, system elements are not always independent; and switching devices are rarely failure-free, and certain parts and components can fail even in standby operation due to the effect of radiation, shock, and internally-generated heat. Thus, the need for a switching device, and the increase in complexity, weight and space as well as interdependence between the failure probabilities of redundant elements may actually reduce the effectiveness of the redundancy application. For a system where the weight or space requirement is of prime importance, the use of redundancy may prove to be a prohibitive design approach even through a redundancy application will improve the desired reliability of the system.

In a certain situation, rather than add a third or fourth parallel redundant element, it may be more economical in terms of weight, space, and cost to increase the reliability of the system by making the system repairable while in operation. If a system is so designed that instant repair work on the failed system element is feasible by an attending operator, reliability improvement can be expected to be much greater than the case of a non-repairable system. Furthermore, this improvement in reliability is achieved without adding much system weight as compared with a duplication of entire system elements.

In actual system applications, whether or not a given system is repairable upon failure is dependent upon the many factors as (1) hardware characteristics, such as physical make-up of the system, complexity, and accessibility of failed elements; (2) availability of spare parts to replace the failed parts; (3) efficiency of failure detecting devices; (4) availability of proper tools to perform the necessary repair work; (5) capability of an attending operator to perform the required repair work on the failed system elements.

Of these factors which will affect the repairability of a failed system element, only the hardware characteristics are considered to be under direct engineering control of the system designers. In order to perform the immediate repair work on a failed element, the system designer must design his system in such a way that failed components or parts can be easily removed and replaced. In this respect, the physical make-up of the system configurations will affect to a large extent the accessibility of the failed system elements. Although the remaining four factors which will either affect repairability of the failed element or increase the difficulties of repair work are not directly under the control of the system designer, he is still responsible for specifying these requirements if he wants to incorporate this type of repair policy as a means of achieving the desired reliability goal.

In developing a reliability model in which system elements are made repairable by an attending operator, it would seem appropriate that some measure should be used to represent the degree of feasibility with which a system element can be repaired. The index of repairability, E, has been selected in the model developed here for this concept. The index of repairability is defined in terms of the conditional probability that an attending operator can repair the failed system element given that a failure has occurred. Thus, it is assumed that in the event of a failure of any system element, the failed element can be repaired by an attending operator with a certain probability.

This probability measure E is somewhat analogous to the current definition of repairability used in reliability literature. However, the difference between these two concepts is that the measure E adopted here is a conditional probability measure which depends upon those factors as hardware characteristics, availability of spare parts and repair tools, efficiency of failure detecting devices, and capability of an attending operator to perform necessary repair work, whereas the repairability is restricted to the probability that a failed system will be restored to operable condition in a specified active repair time.

The actual determination of this factor E is somewhat ambiguous at the early stages of a development program, since the detailed information is not available at this time. In many cases, this factor may be estimated from the available failure data showing the percentage of failures found to be repairable for a similar system with comparable function and complexity. If the objective of the designer is to increase reliability of his system by incorporating this type of repair policy, the system designer must then specify these requirements on his system design so that the desired improvement in reliability can be achieved.

In this chapter, reliability models are developed to determine the reliability functions and the mean-time-to-failures for both non-repairable and repairable system configurations. Then the comparison is made between non-repairable systems and repairable systems in terms of reliability and mean-time-to-failure. Relative improvements in reliability and mean-time-to-failure of repairable systems over its non-repairable systems are estimated to demonstrate the advantages of this repair policy.

The Reliability Function and Mean Time to Failure of Redundant Systems

In order to find the reliability function and mean-time-to-failure of the redundant system under consideration, the following assumptions are made:

1. The system elements considered here are those whose operation can be described in discrete terms as either "operable" or "failed" over a specified time period.

2. The failure time density function of the redundant system is of the exponential form with constant failure rates. Thus, the probability of a single system element failure in $(t, t + dt)$ is $\lambda_1\ dt$ during active operation (on-line), and $\lambda_2\ dt$ while in standby operation (off-line), where λ_1 and λ_2 denote the constant failure rate, respectively.

3. The repair time density function is also of the exponential form with the constant repair rate μ.

4. The probability of more than one failure or repair during $(t, t + dt)$ is assumed to be $O(dt)$, where $O(dt)$ denotes a quantity which is of smaller order of magnitude than dt.

5. The repair work on a failed system element starts immediately by an attending operator given that the failure is properly detected and the spare parts required for the repairs are available. The conditional probability that an attending operator can repair the failed element given that a failure has occurred during the system operation is equal to E.

6. Each element in the redundant system is assumed to be stochastically independent so that the failure or repair in one of the system elements will not affect the failure rate or repair rate of the other system elements.

With the above assumptions, the failure and repair process in a system can be developed in terms of probabilities $P_K(t)$ that at time t the system is in state K. A state of the system is defined simply by listing the conditions of the system elements which are either working satisfactorily or failed. At any instant in time, the system can be in one of a number of mutually exclusive states.

When the failure and repair distributions are the exponential, the probabilities that the system changes from state i to state j satisfy the property of a stationary Markov chain whose state space is the non-negative integers. Let $p_{ij}\ dt$ be the probability that the system changes from state i to state j in time dt. Then the transition probabilities p_{ij} which describe the probabilities of transitions from state i to state j in time dt may be arranged in a matrix of transition probabilities

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ \cdot \\ \cdot \\ \cdot \\ n \end{array} \left[\begin{array}{cccccc} 0 & 1 & 2 & \cdots & n \\ p_{00} & p_{01} & p_{02} & \cdots & p_{0n} \\ p_{10} & p_{11} & p_{12} & \cdots & p_{1n} \\ p_{20} & p_{21} & p_{22} & \cdots & p_{2n} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ p_{n0} & p_{n1} & p_{n2} & \cdots & p_{nn} \end{array} \right]$$

where the subscripts denote the state of the system, and n represents the number of distinct states. Clearly matrix P is a square matrix with non-negative elements and unit row sums.

Given the transition probability matrix, it is possible to derive differential equations for the probabilities $P_K(t)$ of finding the system at time t in state K by the following argument. To calculate $P_K(t + dt)$, first note that at time t + dt the system can be in state K only if one of the following conditions is satisfied: (1) At time t the system is in state K and during (t,t + dt) no transition to the other states occurs; (2) at time t the system is in the state other than state K and during (t,t + dt) a transition to state K occurs;

(3) during $(t,t + dt)$ two or more transitions occur. By the assumption made above, the probability of the last event is $0(dt)$. The first two events are mutually exclusive so that their probabilities add. This means that the functions $P_K(t)$ satisfy the following system of differential equations:

$$P_K(t + dt) = P_K(t)\left\{\begin{array}{l}\text{Probability that no transition from state} \\ \text{K to the other states occurs in} \\ \qquad\qquad (t,t + dt)\end{array}\right\} \tag{5-2}$$

$$+ \sum_{i=K}^{n} P_i(t)\left\{\begin{array}{l}\text{Probability that a transition to} \\ \text{state K occurs in } (t,t + dt)\end{array}\right\}$$

$$= P_K(t)\ [1 - \sum_{i=K}^{n} p_{Ki}\ dt] + \sum_{i=K}^{n} P_i(t) \cdot p_{iK}\ dt$$

Transposing the term $P_K(t)$ and dividing the equation by dt, and taking the limit as $dt \to 0$, one obtains

$$P_K'(t) = \sum_{i=K}^{n} P_i(t) \cdot p_{iK} - \sum_{i=K}^{n} P_K(t) \cdot p_{Ki} \tag{5-3}$$

where $K = 0,1,2,\ldots,n$.

Given the initial probability distribution $\{P_K(0)\}$ for $K = 0,1,2,\ldots,n$, one can solve (5-3) to obtain $P_K(t)$. In the following section, the reliability function which gives the probability that the system will not have reached a failed state within a time interval $(0,t)$, and the mean-time-to-failure are derived for both non-repairable and repairable redundant systems.

Non-Repairable Redundant System with an Imperfect Switching Device:

Consider a standby redundant system with elements A and B and a switching device S as shown in Figure 11.

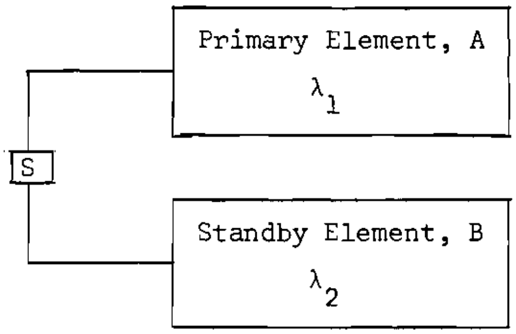


Figure 11. Two-Element Redundant System with an Imperfect Switching Device

Assume that the system element has a failure rate λ_1 while in on-line, and λ_2 while in off-line. The switching device may fail any time and has a failure rate λ_S. Therefore, if the switch S fails, standby element cannot be used even at a time of the primary element failure. To derive the reliability function and mean-time-to-failure for this case, one defines the following eight possible states of the system, where the states denote the different conditions of the system elements, and employs the following transition probability matrix P. The bar over the letter represents failure of that element. The first three states $\bar{A}\ \bar{S}\ \bar{B}$, $\bar{A}\ S\ \bar{B}$, and $\bar{A}\ \bar{S}\ B$ are considered unacceptable states (failed states). Thus, when the system once reaches these states no transitions out of these states are possible, and the transition process is terminated. The remaining states constitute the acceptable states, and transitions between these states are possible with the specified transition probabilities.

$$
P = \begin{array}{r} \\ (1)\bar{A}\ \bar{S}\ \bar{B} \\ (2)\bar{A}\ S\ \bar{B} \\ (3)\bar{A}\ \bar{S}\ B \\ (4)\bar{A}\ S\ B \\ (5)A\ \bar{S}\ \bar{B} \\ (6)A\ S\ \bar{B} \\ (7)A\ \bar{S}\ B \\ (8)A\ S\ B \end{array}
\begin{array}{c} \begin{array}{cccccccc} \bar{A}\,\bar{S}\,\bar{B} & \bar{A}\,S\,\bar{B} & \bar{A}\,\bar{S}\,B & \bar{A}\,S\,B & A\,\bar{S}\,\bar{B} & A\,S\,\bar{B} & A\,\bar{S}\,B & A\,S\,B \end{array} \\
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & \lambda_1 & \lambda_s & 1-(\lambda_1+\lambda_s) & 0 & 0 & 0 & 0 \\
\lambda_1 & 0 & 0 & 0 & 1-\lambda_1 & 0 & 0 & 0 \\
0 & \lambda_1 & 0 & 0 & \lambda_s & 1-(\lambda_1+\lambda_s) & 0 & 0 \\
0 & 0 & \lambda_1 & 0 & \lambda_2 & 0 & 1-(\lambda_1+\lambda_2) & 0 \\
0 & 0 & 0 & \lambda_1 & 0 & \lambda_2 & \lambda_s & 1-(\lambda_1+\lambda_2+\lambda_s)
\end{bmatrix} \end{array}
$$

(5-4)

The foregoing transition probabilities then lead to the following probability equations in terms of $P_K(t)$ (see Equation (5-2)):

$$P_1(t + dt) = P_1(t) + \lambda_1 P_5(t)\,dt$$

$$P_2(t + dt) = P_2(t) + \lambda_1 P_4(t)\,dt + \lambda_1 P_6(t)\,dt$$

$$P_3(t + dt) = P_3(t) + \lambda_s P_4(t)\,dt + \lambda_1 P_7(t)\,dt$$

$$P_4(t + dt) = P_4(t)\,[1 - (\lambda_1 + \lambda_s)\,dt] + \lambda_1 P_8(t)\,dt \tag{5-5}$$

$$P_5(t + dt) = P_5(t)\,[1 - \lambda_1\,dt] + \lambda_s P_6(t)\,dt + \lambda_2 P_7(t)\,dt$$

$$P_6(t + dt) = P_6(t)\,[1 - (\lambda_1 + \lambda_s)\,dt] + \lambda_2 P_8(t)\,dt$$

$$P_7(t + dt) = P_7(t)\,[1 - (\lambda_1 + \lambda_2)dt] + \lambda_s P_8(t)\,dt$$

$$P_8(t + dt) = P_8(t)\,[1 - (\lambda_1 + \lambda_2 + \lambda_s)\,dt]$$

Transposing the term $P_K(t)$ for $K = 1,2,\ldots,8$, and dividing the equations by dt, and letting $dt \to 0$, one obtains the following system of differential equations:

$$P_1'(t) - \lambda_1 P_5(t) = 0 \tag{5-6}$$

$$P_2'(t) - \lambda_1 P_4(t) - \lambda_1 P_6(t) = 0$$

$$P_3'(t) - \lambda_s P_4(t) - \lambda_1 P_7(t) = 0$$

$$P_4'(t) + (\lambda_1 + \lambda_s)\,P_4(t) - \lambda_1 P_8(t) = 0$$

$$P_5'(t) + \lambda_1 P_5(t) - \lambda_s P_6(t) - \lambda_2 P_7(t) = 0$$

$$P_6'(t) + (\lambda_1 + \lambda_s) P_6(t) - \lambda_2 P_8(t) = 0$$

$$P_7'(t) + (\lambda_1 + \lambda_2) P_7(t) - \lambda_s P_8(t) = 0$$

$$P_8'(t) + (\lambda_1 + \lambda_2 + \lambda_s) P_8(t) = 0$$

Assuming that at time $t = 0$, both system elements and switching devices are operable, the initial conditions are then given by:

$$P_1(0) = P_2(0) = \ldots = P_7(0) = 0, \text{ and } P_8(0) = 1 \tag{5-7}$$

In order to solve (5-6) for $P_K(t)$, $K = 1,2,\ldots,8$, one employs the Laplace transforms of (5-6), which yield the following simultaneous linear equations with the constant coefficients:

$$SP_1(s) - \lambda_1 P_5(s) = 0 \tag{5-8}$$

$$SP_2(s) - \lambda_1 P_4(s) - \lambda_1 P_6(s) = 0$$

$$SP_3(s) - \lambda_s P_4(s) - \lambda_1 P_7(s) = 0$$

$$SP_4(s) + (\lambda_1 + \lambda_s) P_4(s) - \lambda_1 P_8(s) = 0$$

$$SP_5(s) + \lambda_1 P_5(s) - \lambda_s P_6(s) - \lambda_2 P_7(s) = 0$$

$$SP_6(s) + (\lambda_1 + \lambda_s) P_6(s) - \lambda_2 P_8(s) = 0$$

$$SP_7(s) + (\lambda_1 + \lambda_2) P_7(s) - \lambda_s P_8(s) = 0$$

$$SP_8(s) + (\lambda_1 + \lambda_2 + \lambda_s)\ P_8(s) = 1$$

Since the system is in operable state if the system is in one of states 4, 5, 6, 7, and 8, one needs only to solve $P_K(s)$ for these states to obtain the reliability function of this system.

Solving for $P_K(s)$, $K = 4,5,6,7,8$, one obtains:

$$P_8(s) = \frac{1}{S + \lambda_1 + \lambda_2 + \lambda_s}$$

$$P_7(s) = \frac{\lambda_s}{(S + \lambda_1 + \lambda_2)(S + \lambda_1 + \lambda_2 + \lambda_s)}$$

$$P_6(s) = \frac{\lambda_2}{(S + \lambda_1 + \lambda_s)(S + \lambda_1 + \lambda_2 + \lambda_s)}$$

$$P_5(s) = \frac{\lambda_2\ \lambda_s}{(S + \lambda_1)(S + \lambda_1 + \lambda_s)(S + \lambda_1 + \lambda_2 + \lambda_s)}$$

$$+ \frac{\lambda_2\ \lambda_s}{(S + \lambda_1)(S + \lambda_1 + \lambda_2)(S + \lambda_1 + \lambda_2 + \lambda_s)}$$

$$P_4(s) = \frac{\lambda_1}{(S + \lambda_1 + \lambda_s)(S + \lambda_1 + \lambda_2 + \lambda_s)}$$

Taking the Laplace inverse transforms for $P_K(s)$, $K = 4,5,6,7,8$, one finds:

$$P_4(t) = \frac{\lambda_1}{\lambda_2}\left[e^{-(\lambda_1+\lambda_s)t} - e^{-(\lambda_1+\lambda_2+\lambda_s)t}\right]$$

$$P_5(t) = e^{-\lambda_1 t} - e^{-(\lambda_1+\lambda_s)t} - e^{-(\lambda_1+\lambda_2)t} + e^{-(\lambda_1+\lambda_2+\lambda_s)t}$$

$$P_6(t) = e^{-(\lambda_1+\lambda_s)t} - e^{-(\lambda_1+\lambda_2+\lambda_s)t}$$

$$P_7(t) = e^{-(\lambda_1+\lambda_2)t} - e^{-(\lambda_1+\lambda_2+\lambda_s)t}$$

$$P_8(t) = e^{-(\lambda_1+\lambda_2+\lambda_s)t}$$

Therefore, the reliability function for this redundant system is found to be

$$R(t) = P_4(t) + P_5(t) + P_6(t) + P_7(t) + P_8(t) \tag{5-9}$$

$$= e^{-\lambda_1 t} + \frac{\lambda_1}{\lambda_2} e^{-(\lambda_1+\lambda_s)t} - \frac{\lambda_1}{\lambda_2} e^{-(\lambda_1+\lambda_2+\lambda_s)t}$$

The mean-time-to-failure (MTTF) of this system is obtained by integrating (5-9) over its range $(0,\infty)$.

$$\text{MTTF} = \int_0^\infty R(t)dt = \frac{1}{\lambda_1} + \frac{\lambda_1}{\lambda_2(\lambda_1+\lambda_s)} - \frac{\lambda_1}{\lambda_2(\lambda_1+\lambda_2+\lambda_s)} \tag{5-10}$$

It is interesting to note that when one has a perfect switching device, i.e., $\lambda_s = 0$, one obtains from (5-9) and (5-10) that

$$R(t) = \frac{\lambda_1 + \lambda_2}{\lambda_2} e^{-\lambda_1 t} - \frac{\lambda_1}{\lambda_2} e^{-(\lambda_1+\lambda_2)t} \tag{5-11}$$

$$MTTF = \frac{1}{\lambda_1} + \frac{1}{\lambda_1 + \lambda_2} \tag{5-12}$$

When $\lambda_1 = \lambda_2 = \lambda$, which is the case of an active redundancy, one has

$$R(t) = 2e^{-\lambda t} - e^{-2\lambda t} \tag{5-13}$$

$$MTTF = \frac{3}{2\lambda} \tag{5-14}$$

Non-Repairable Standby Redundant System

Consider the case of a doubly redundant system shown in Figure 12.

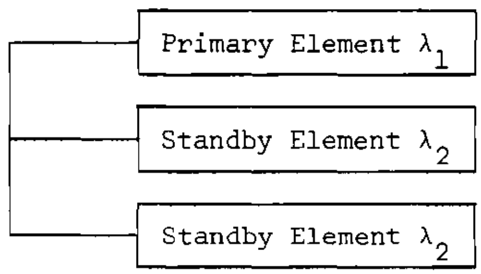


Figure 12. Three-Element Standby Redundant System Configuration

Again assume that the failure rate of an on-line element is λ_1 and that of off-line element is λ_2. At any one time the system is defined to be operable if at least one of the three system elements is in the operable state. The transition probability matrix P, where the state space is 0,1,2, and 3, representing the number of system elements operable, is given by

$$P = \begin{array}{c c} & \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \end{array} & \begin{bmatrix} 1 & 0 & 0 & 0 \\ \lambda_1 & 1-\lambda_1 & 0 & 0 \\ 0 & \lambda_1+\lambda_2 & 1-(\lambda_1+\lambda_2) & 0 \\ 0 & 0 & \lambda_1+2\lambda_2 & 1-(\lambda_1+2\lambda_2) \end{bmatrix} \end{array} \qquad (5\text{-}15)$$

From the above transition probability matrix and Equation (5-3), one obtains the following system of differential equations:

$$P_0'(t) - \lambda_1 P_1(t) = 0$$

$$P_1'(t) + \lambda_1 P_1(t) - (\lambda_1 + \lambda_2) P_2(t) = 0$$

$$P_2'(t) + (\lambda_1 + \lambda_2) P_2(t) - (\lambda_1 + 2\lambda_2) P_3(t) = 0 \qquad (5\text{-}16)$$

$$P_3'(t) + (\lambda_1 + 2\lambda_2) P_3(t) = 0$$

with the initial conditions

$$P_0(0) = P_1(0) = P_2(0) = 0, \text{ and } P_3(0) = 1 \tag{5-17}$$

Taking the Laplace transforms of (5-16) gives

$$S\, P_0(s) - \lambda_1\, P_1(s) = 0$$

$$S\, P_1(s) + \lambda_1\, P_1(s) - (\lambda_1 + \lambda_2)\, P_2(s) = 0$$

$$S\, P_2(s) + (\lambda_1 + \lambda_2)\, P_2(s) - (\lambda_1 + 2\lambda_2)\, P_3(s) = 0$$

$$S\, P_3(s) + (\lambda_1 + 2\lambda_2)\, P_3(s) = 1$$

Solving for $P_K(s)$, K = 1, 2, 3, and taking the Laplace inverse transforms of $P_K(s)$, one finds:

$$P_1(t) = (\lambda_1 + \lambda_2)\,(\lambda_1 + 2\lambda_2)\left[\frac{1}{2\lambda_2^2}\, e^{-\lambda_1 t} - \frac{1}{\lambda_2^2}\, e^{-(\lambda_1 + \lambda_2)t} + \frac{1}{2\lambda_2^2}\, e^{-(\lambda_1 + 2\lambda_2)t}\right]$$

$$P_2(5) = \frac{\lambda_1 + 2\lambda_2}{\lambda_2}\left[e^{-(\lambda_1+\lambda_2)t} - e^{-(\lambda_1 + 2\lambda_2)t}\right]$$

$$P_3(t) = e^{-(\lambda_1+2\lambda_2)t}$$

Therefore, the reliability function for this redundant system is found to be

$$R(t) = P_1(t) + P_2(t) + P_3(t) \tag{5-18}$$

$$= \left[\frac{(\lambda_1 + \lambda_2)\,(\lambda_1 + 2\lambda_2)}{2\lambda_2^2}\right] e^{-\lambda_1 t}$$

$$+ \left[\frac{\lambda_2(\lambda_1 + 2\lambda_2) - (\lambda_1 + \lambda_2)\,(\lambda_1 + 2\lambda_2)}{\lambda_2^2}\right] e^{-(\lambda_1 + \lambda_2)t}$$

$$+ \left[\frac{(\lambda_1 + \lambda_2)\,(\lambda_1 + 2\lambda_2) - 2\lambda_2(\lambda_1 + 2\lambda_2) + 2\lambda_2^2}{2\lambda_2^2}\right] e^{-(\lambda_1 + 2\lambda_2)t}$$

The reliability function for a three-element active redundant system is obtained from (5-18) by letting $\lambda_1 = \lambda_2 = \lambda$ as

$$R(t) = 3\, e^{-\lambda t} - 3\, e^{-2\lambda t} + e^{-3\lambda t} \tag{5-19}$$

The mean-time-to-failure of three-element standby redundant system may be obtained as before by integrating the reliability function in (5-18) over its range $(0,\infty)$. However, a simpler means of obtaining MTTF is the use of transition matrix P. The matrix in (5-15) is the transition matrix of an absorbing Markov chain with state 0 as the absorbing state (the state of complete system failure). Once the absorbing state is reached, it stays there with unit probability since this, by definition, terminates the transition process of the system.

Any transition probability matrix of an absorbing Markov chain with s transient states and r - s absorbing states can be expressed in the following canonical form

$$P = \begin{array}{cc} & \begin{array}{cc} r\text{-}s & s \end{array} \\ & \left[\begin{array}{c|c} I & 0 \\ \hline R & Q \end{array}\right] \begin{array}{c} r\text{-}s \\ s \end{array} \end{array}$$

where I is the r-s by r-s identity matrix, and 0 is the r-s by s zero matrix.

For an absorbing Markov chain, Kemeny and Snell define the fundamental matrix as[1]

$$N = [I - Q]^{-1}$$

If one defines N_{ij} to be the total number of times the system is in state j before absorption if it starts in state i and

$$E[N_{ij}] = n_{ij}$$

to be the expected number of visits to state j before absorption,

---

[1]John G. Kemeny and J. Laurie Snell, *Finite Markov Chains* (Princeton, N. J.: D. Van Nostrand Company, Inc., 1960), p. 46.

starting in state i, then it can be shown that[2]

$$N = [I - Q]^{-1} = [n_{ij}]$$

Using the above property of the absorbing Markov Chain, the transition probability matrix in Equation (5-15) can be written in canonical form as

$$P = \begin{array}{c c} & \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \end{array} & \left[\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline \lambda_1 & 1-\lambda_1 & 0 & 0 \\ 0 & \lambda_1+\lambda_2 & 1-(\lambda_1+\lambda_2) & 0 \\ 0 & 0 & \lambda_1+2\lambda_2 & 1-(\lambda_1+2\lambda_2) \end{array}\right] \end{array}$$

and

$$I - Q = \begin{array}{c c} & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \left[\begin{array}{ccc} \lambda_1 & 0 & 0 \\ -\lambda_1-\lambda_2 & \lambda_1+\lambda_2 & 0 \\ 0 & -\lambda_1-2\lambda_2 & \lambda_1+2\lambda_2 \end{array}\right] \end{array}$$

---

[2] *Ibid.*, p. 47.

Upon inverting the above I - Q matrix, one has

$$N = \frac{1}{\lambda_1(\lambda_1+\lambda_2)(\lambda_1+2\lambda_2)} \begin{bmatrix} (\lambda_1+\lambda_2) & (\lambda_1+2\lambda_2) & 0 & 0 \\ (\lambda_1+\lambda_2) & (\lambda_1+2\lambda_2) & \lambda_1(\lambda_1+2\lambda_2) & 0 \\ (\lambda_1+\lambda_2) & (\lambda_1+2\lambda_2) & \lambda_1(\lambda_1+2\lambda_2) & \lambda_1(\lambda_1+\lambda_2) \end{bmatrix}$$

It is seen from the above N matrix that the system spends

$(\lambda_1+\lambda_2)(\lambda_1+2\lambda_2) \,/\, \lambda_1(\lambda_1+\lambda_2)(\lambda_1+2\lambda_2)$ times in state 1, and

$\lambda_1(\lambda_1+2\lambda_2) \,/\, \lambda_1(\lambda_1+\lambda_2)(\lambda_1+2\lambda_2)$ times in state 2, and

$\lambda_1(\lambda_1+\lambda_2) \,/\, \lambda_1(\lambda_1+\lambda_2)(\lambda_1+2\lambda_2)$ times in state 3 before

absorption in state 0 (complete system failure), given that the system initially starts with state 3 (all system elements operable). The sum of these quantities is, by definition, the mean-time-to-failure of this redundant system. Thus,

$$\mathrm{MTTF} = \frac{(\lambda_1+\lambda_2)\ (\lambda_1+2\lambda_2) + \lambda_1(\lambda_1+2\lambda_2) + \lambda_1(\lambda_1+\lambda_2)}{\lambda_1(\lambda_1+\lambda_2)\ (\lambda_1+2\lambda_2)} \qquad (5\text{-}20)$$

$$= \frac{1}{\lambda_1} + \frac{1}{\lambda_1 + \lambda_2} + \frac{1}{\lambda_1 + 2\lambda_2}$$

For a n-element non-repairable standby redundant system, the transition matrix P and I - Q matrix are respectively

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ \cdot \\ \cdot \\ \cdot \\ n-1 \\ n \end{array} \begin{array}{c} \begin{array}{cccccccc} 0 & 1 & 2 & \cdot & \cdot & \cdot & n-1 & n \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & \cdot & \cdot & \cdot & 0 & 0 \\ \lambda_1 & 1-\lambda_1 & 0 & \cdot & \cdot & \cdot & 0 & 0 \\ 0 & \lambda_1+\lambda_2 & 1-(\lambda_1+\lambda_2) & \cdot & \cdot & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & 0 & 1-[\lambda_1+(n-2)\lambda_2] & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_1+(n-1)\lambda_2 & 1-[\lambda_1+(n-1)\lambda_2] \end{bmatrix} \end{array}$$

and

$$I-Q = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ \cdot \\ \cdot \\ \cdot \\ n-1 \\ n \end{array} \begin{array}{c} \begin{array}{cccccccc} 1 & 2 & 3 & \cdot & \cdot & \cdot & n-1 & n \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & \cdot & \cdot & \cdot & 0 & 0 \\ -(\lambda_1+\lambda_2) & \lambda_1+\lambda_2 & 0 & \cdot & \cdot & \cdot & 0 & 0 \\ 0 & -(\lambda_1+2\lambda_2) & \lambda_1+2\lambda_2 & \cdot & \cdot & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & \lambda_1+(n-2)\lambda_2 & 0 \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & -[\lambda_1+(n-1)\lambda_2] & \lambda_1+(n-1)\lambda_2 \end{bmatrix} \end{array}$$

Upon inverting the I - Q matrix, one finds the mean-time-to-failure of this n-element redundant system as

$$MTTF = \frac{1}{\lambda_1} + \frac{1}{\lambda_1 + \lambda_2} + \frac{1}{\lambda_1 + 2\lambda_2} + \cdots + \frac{1}{\lambda_1 + (n-1)\lambda_2} \qquad (5\text{-}21)$$

$$= \sum_{K=0}^{n-1} \frac{1}{(\lambda_1 + K\lambda_2)}$$

For a n-element parallel redundant system where $\lambda_1 = \lambda_2 = \lambda$, the mean-time-to-failure is obtained from Equation (5-21) as

$$MTTF = \frac{1}{\lambda} + \frac{1}{2\lambda} + \frac{1}{3\lambda} + \cdots + \frac{1}{n\lambda} \qquad (5\text{-}22)$$

$$= \sum_{K=0}^{n} \frac{1}{K\lambda}$$

The comparison of Equation (5-21) with (5-22) reveals that when the failure rate of off-line elements is less than that of on-line elements which is true in actual situations, standby redundancy has greater mean-time-to-failure than active redundancy provided that the failure rate of a switching device is negligible.

Repairable Standby Redundant System, Type I

Figure 13 shows a standby redundant system where only a primary system element is made repairable.

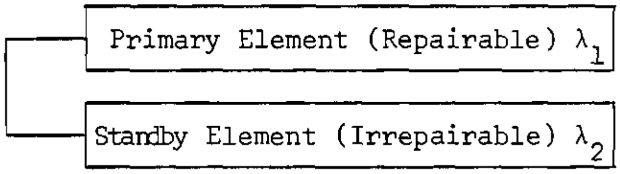


Figure 13. Repairable Standby Redundant System, Type I

The standby element is activated while the repair work is being done on the failed primary element, and is placed again in standby position as soon as the repair work is completed on the failed primary element. At any instant in time, the system can be in one of the following four mutually exclusive states:

State 0: None of the two system elements are operable

State 1: One element is operable, the other failed, but irrepairable

State 2: One element is operable, the other failed and repair work is being performed by an attending operator

State 3: Both elements are operable

The system makes transitions from states 1, 2, and 3. However, no transition is possible from state 0 to any other state. Once the state 0 is reached, the system stays there with unit probability. Then the transition probability matrix for this redundant system is

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ \lambda_1 & 1-\lambda_1 & 0 & 0 \\ \lambda_1 & 0 & 1-(\lambda_1+\mu) & \mu \\ 0 & (1-E)\lambda_1+\lambda_2 & E\lambda_1 & 1-(\lambda_1+\lambda_2) \end{bmatrix} \end{array} \tag{5-23}$$

where E denotes the conditional probability that an attending operator can repair the failed system given that a failure has occurred.

Proceeding as previously one can find the resulting system of differential equations as:

$$P_0'(t) - \lambda_1 P_1(t) - \lambda_1 P_2(t) = 0$$

$$P_1'(t) + \lambda_1 P_1(t) - [(1-E)\lambda_1 + \lambda_2]P_3(t) = 0$$

$$P_2'(t) + (\lambda_1 + \mu) P_2(t) - E\lambda_1 P_3(t) = 0$$

$$P_3'(t) - \mu P_2(t) + (\lambda_1 + \lambda_2) P_3(t) = 0$$

with initial conditions

$$P_0(0) = P_1(0) = P_2(0) = 0, \text{ and } P_3(0) = 1$$

The reliability function to be found is of the form $R(t) = P_1(t) + P_2(t) + P_3(t)$, since states 1, 2, and 3 are the acceptable states. In order to find $P_K(t)$, $K = 1,2,3$, one takes the Laplace transforms of the foregoing differential equations and obtains the following linear equations:

$$S\ P_0(s) - \lambda_1\ P_1(s) - \lambda_1\ P_2(s) = 0$$

$$S\ P_1(s) + \lambda_1\ P_1(s) - [(1-E)\lambda_1 + \lambda_2]P_3(s) = 0$$

$$S\ P_2(s) + (\lambda_1 + \mu)\ P_2(s) - E\lambda_1\ P_3(s) = 0 \tag{5-24}$$

$$S\ P_3(s) - \mu P_2(s) + (\lambda_1 + \lambda_2)\ P_3(s) = 1$$

Solving first for $P_2(s)$ from the last two equations one has

$$P_2(s) = \frac{E\lambda_1}{S^2 + (2\lambda_1 + \mu + \lambda_2)S + \lambda_1(\lambda_1 + \mu + \lambda_2 + E\mu) + \mu\lambda_2}$$

Solving for the roots of denominator

$$S^2 + (2\lambda_1 + \mu + \lambda_2)S + \lambda_1(\lambda_1 + \mu + \lambda_2 + E\mu) + \mu\lambda_2$$

One obtains two roots as

$$S_1 = \frac{-(2\lambda_1 + \mu + \lambda_2) + \sqrt{(\mu - \lambda_2)^2 - 4\ E\mu\ \lambda_1}}{2} \tag{5-25}$$

$$S_2 = \frac{-(2\lambda_1 + \mu + \lambda_2) - \sqrt{(\mu - \lambda_2)^2 - 4\,E\mu\,\lambda_1}}{2}$$

Applying the method of partial fractions, $P_2(s)$ can now be written as

$$P_2(s) = \frac{E\,\lambda_1}{S_1 - S_2}\left[\frac{1}{S - S_1} - \frac{1}{S - S_2}\right]$$

Substituting $P_2(s)$ found above into the last equation of (5-24), one has

$$P_3(s) = \frac{S + \lambda_1 + \mu}{E\lambda_1}\;P_2(s) = \frac{S + \lambda_1 + \mu}{(S - S_1)\,(S - S_2)}$$

and from the second equation of (5-24)

$$P_1(s) = [(1 - E)\lambda_1 + \lambda_2]\,\Big(\frac{1}{(S - S_1)(S - S_2)} + \frac{\mu}{(S_1 + \lambda_1)(S_2 + \lambda_2)(S + \lambda_1)}$$

$$+ \frac{\mu}{(S_1 + \lambda_1)(S_1 - S_2)(S - S_1)} - \frac{\mu}{(S_2 + \lambda_1)(S_1 - S_2)(S - S_2)}\Big)$$

Taking the Laplace inverse transforms of $P_1(s)$, $P_2(s)$ and $P_3(s)$ found above, one finds:

$$P_1(t) = [(1\text{-}E)\lambda_1 + \lambda_2]\,\Big(\frac{1}{S_1 - S_2}\,(e^{S_1 t} - e^{S_2 t})$$

$$+ \frac{\mu}{(S_1 + \lambda_1)(S_2 + \lambda_1)} e^{-\lambda_1 t} + \frac{\mu}{(S_1 + \lambda_1)(S_1 - S_2)} e^{S_1 t}$$

$$- \frac{\mu}{(S_2 + \lambda_1)(S_1 - S_2)} e^{S_2 t}]$$

$$P_2(t) = \frac{E\ \lambda_1}{S_1 - S_2} \left(e^{S_1 t} - e^{S_2 t}\right)$$

$$P_3(t) = \frac{S_1 + \mu + \lambda_1}{S_1 - S_2} e^{S_1 t} - \frac{S_2 + \lambda_1 + \mu}{S_1 - S_2} e^{S_2 t}$$

Therefore, the reliability function for this redundant system where only a primary element is repairable with probability E is given by

$$R(t) = P_1(t) + P_2(t) + P_3(t) \tag{5-26}$$

$$= \frac{[(1 - E)\lambda_1 + \lambda_2](1 + \mu) + E\lambda_1 + S_1 + \mu + \lambda_1}{S_1 - S_2} e^{S_1 t}$$

$$- \frac{[(1 - E)\lambda_1 + \lambda_2](1 + \mu) + E\lambda_1 + S_2 + \mu + \lambda_1}{S_1 - S_2} e^{S_2 t}$$

$$+ \frac{[(1 - E)\lambda_1 + \lambda_2]\mu}{(S_1 + \lambda_1)(S_2 + \lambda_1)} e^{-\lambda_1 t}$$

The reliability function in Equation (5-26) reduces to the same reliability function for a non-repairable standby system found in Equation (5-11) when the primary element is not made repairable, i.e., $\mu = 0$, and $E = 0$. In order to find the mean-time-to-failure of this system, one makes use of the transition probability matrix P in Equation (5-23). The fundamental matrix is found to be

$$I - Q = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1+\mu & -\mu \\ -(1-E)\lambda_1-\lambda_2 & -E\lambda_1 & \lambda_1+\lambda_2 \end{bmatrix} \end{array}$$

Upon inverting I - Q matrix, one finds

$$N = \frac{1}{\lambda_1(\lambda_1+\lambda_2)(\lambda_1+\mu)-\mu E\lambda_1^2} \begin{bmatrix} (\lambda_1+\mu)(\lambda_1+\lambda_2)-E\mu\lambda_1 & 0 & 0 \\ [(1-E)\lambda_1+\lambda_2]\mu & \lambda_1(\lambda_1+\mu) & \mu\lambda_1 \\ (\lambda_1+\mu)[(1-E)\lambda_1+\lambda_2] & E\lambda_1^2 & \lambda_1(\lambda_1+\mu) \end{bmatrix}$$

Then the mean-time-to-failure is given by

$$MTTF = \frac{(\lambda_1 + \mu)\ [(2 - E)\lambda_1 + \lambda_2] + E\lambda_1^2}{\lambda_1(\lambda_1 + \lambda_2)(\lambda_1 + \mu) - E\mu\lambda_1^2} \qquad (5\text{-}27)$$

It is interesting to note that when $\mu = 0$, and $E = 0$, Equations (5-26) and (5-27) reduce to the equations found for non-repairable standby system in Equations (5-11) and (5-12).

Repairable Standby Redundant System, Type II

Figure 14 shows a two-element standby system where both system elements are made repairable with the probability E for each element.

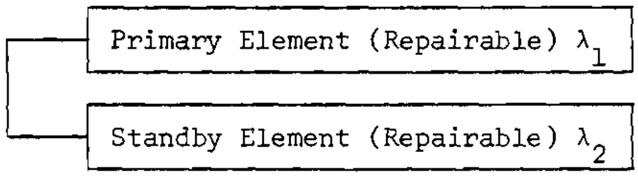


Figure 14. Repairable Standby Redundant System, Type II

The transition probability matrix with state 0 as the absorbing state (complete failure) is then

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \!\!\! \begin{array}{c} \begin{array}{cccc} 0 & \quad 1 & \quad\quad 2 & \quad\quad 3 \end{array} \\ \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ \lambda_1 & 1-\lambda_1 & 0 & 0 \\ \lambda_1 & 0 & 1-(\lambda_1+\mu) & \mu \\ 0 & (1-E)(\lambda_1+\lambda_2) & E(\lambda_1+\lambda_2) & 1-(\lambda_1+\lambda_2) \end{array}\right] \end{array}$$

Then the resulting system of differential equations for this system becomes:

$$P_0'(t) - \lambda_1 P_1(t) - \lambda_1 P_2(t) = 0$$

$$P_1'(t) + \lambda_1 P_1(t) - (1 - E)(\lambda_1+\lambda_2) P_3(t) = 0$$

$$P_2'(t) + (\lambda_1+\mu)\, P_2(t) - E(\lambda_1+\lambda_2)\, P_3(t) = 0$$

$$P_3'(t) - \mu\, P_2(t) + (\lambda_1+\lambda_2)\, P_3(t) = 0$$

with initial conditions

$$P_0(0) = P_1(0) = P_2(0) = 0, \quad \text{and} \quad P_3(0) = 1$$

The reliability function to be found is of the form $R(t) = P_1(t) + P_2(t) + P_3(t)$, since state 1, 2, and 3 are the acceptable states. In order to find $P_K(t)$, $K = 1, 2, 3$, one proceeds as before to take the Laplace transforms of the above differential equations, and then taking the inverse transforms one finds the reliability function as

$$R(t) = P_1(t) + P_2(t) + P_3(t)$$

$$= \left[\frac{S_1 + 2\lambda_1 + \lambda_2 + \mu}{S_1 - S_2} + \frac{(1 - E)\,(\lambda_1 + \lambda_2)\mu}{(S_1 + \lambda_1)(S_1 - S_2)}\right] e^{S_1 t} \tag{5-29}$$

$$- \left[\frac{S_2 + 2\lambda_1 + \lambda_2 + \mu}{S_1 - S_2} + \frac{(1 - E)(\lambda_1+\lambda_2)\mu}{(S_2+\lambda_1)(S_1-S_2)}\right] e^{S_2 t}$$

$$+ \left[\frac{(1 - E)(\lambda_1+\lambda_2)\mu}{(S_1+\lambda_1)(S_2+\lambda_1)}\right] e^{-\lambda_1 t}$$

where

$$S_1 = \frac{-(2\lambda_1 + \lambda_2 + \mu) + \sqrt{(\mu - \lambda_2)^2 + 4\,E\mu(\lambda_1 + \lambda_2)}}{2}$$

$$S_2 = \frac{-(2\lambda_1 + \lambda_2 + \mu) - \sqrt{(\mu - \lambda_2)^2 + 4\,E\mu(\lambda_1 + \lambda_2)}}{2}$$

To find the mean-time-to-failure of this system, one finds the following fundamental matrix

$$I - Q = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1+\mu & -\mu \\ -(1-E)(\lambda_1+\lambda_2) & -E(\lambda_1+\lambda_2) & \lambda_1+\lambda_2 \end{bmatrix} \end{array}$$

Upon inverting I - Q matrix, one finds:

$$N = \frac{1}{\lambda_1(\lambda_1+\mu)(\lambda_1+\lambda_2) - E\mu\lambda_1(\lambda_1+\lambda_2)} \begin{bmatrix} (\lambda_1+\mu)(\lambda_1+\lambda_2)-E\mu(\lambda_1+\lambda_2) & 0 & 0 \\ \mu(1-E)(\lambda_1+\lambda_2) & \lambda_1(\lambda_1+\lambda_2) & \mu\lambda_1 \\ (1-E)(\lambda_1+\mu)(\lambda_1+\lambda_2) & E\lambda_1(\lambda_1+\lambda_2) & \lambda_1(\lambda_1+\mu) \end{bmatrix}$$

Therefore, the mean-time-to-failure of this system is obtained by summing the last row of N matrix as

$$\text{MTTF} = \frac{(\lambda_1 + \lambda_2)\,[\lambda_1 + (1-E)\mu] + \lambda_1(\lambda_1 + \mu)}{\lambda_1(\lambda_1 + \mu)(\lambda_1 + \lambda_2) - E\mu\lambda_1(\lambda_1 + \lambda_2)} \tag{5-30}$$

Again, Equation (5-29) and (5-30) reduce to its non-repairable counterparts of Equations (5-11) and (5-12) when $\mu = 0$, and $E = 0$.

Relative Improvement in Reliability

In order to evaluate the reliability gains obtained by making a redundant system repairable while in operation, the following measures of relative improvement are defined:

$$\begin{matrix}\text{Relative Improvement} \\ \text{in Reliability}\end{matrix} = \frac{\left\{\begin{matrix}\text{Reliability of} \\ \text{Repairable} \\ \text{System}\end{matrix}\right\} - \left\{\begin{matrix}\text{Reliability of} \\ \text{Non-Repairable} \\ \text{System}\end{matrix}\right\}}{\text{Reliability of Non-Repairable System}}$$

$$\begin{matrix}\text{Relative Improvement in} \\ \text{Mean-Time-to-Failure}\end{matrix} = \frac{\left\{\begin{matrix}\text{Mean-Time-to-Failure} \\ \text{of} \\ \text{Repairable System}\end{matrix}\right\}}{\left\{\begin{matrix}\text{Mean-Time-to-Failure} \\ \text{of} \\ \text{Non-Repairable System}\end{matrix}\right\}}$$

Table 11 shows the numerical results for this reliability improvement; it gives the relative improvement in reliability of a two-element repairable system for several values of λ_1 and E. Table 12 then gives the relative improvement in a repairable system's MTTF to a non-

Table 11. Relative Improvement in Reliability of a Two-Element Redundant System, Where $\mu = 1.0$, $\lambda_2 = \lambda_1/2$, $t = 100$

	FAILURE RATE OF ON-LINE ELEMENT, λ_1													
	0.001		0.002		0.004		0.006		0.008		0.01		0.04	
E	R_2	ΔR	R_2	ΔR	R_2	ΔR	R_2	ΔR	R_2	ΔR	R_2	ΔR	R_2	ΔR
0.1	0.9937	0.07	0.9768	0.23	0.9204	0.77	0.8456	1.48	0.7625	2.27	0.6779	3.12	0.0576	15.3
0.2	0.9944	0.13	0.9792	0.47	0.9277	1.57	0.8586	3.04	0.7808	4.72	0.7000	6.55	0.0684	36.7
0.4	0.9957	0.26	0.9839	0.96	0.9432	3.27	0.8869	6.44	0.8217	10.2	0.7524	14.5	0.1059	111.8
0.6	0.9971	0.40	0.9889	1.47	0.9599	5.11	0.9188	10.3	0.8695	16.6	0.8154	24.1	0.1873	274.8
0.8	0.9985	0.54	0.9941	2.00	0.9781	7.09	0.9545	14.6	0.9253	24.1	0.8921	35.7	0.3723	644.7
1.0	0.9999	0.68	0.9994	2.55	0.9977	9.23	0.9948	19.4	0.9907	32.9	0.9856	49.9	0.8052	1510.1
R_1	0.9931		0.9746		0.9133		0.8333		0.7456		0.6574		0.0499	

NOTE: R_1 = Reliability of Two-Element Non-Repairable System ($\mu = 0$)

R_2 = Reliability of Two-Element Repairable System ($\mu = 1.0$)

$$\Delta R = \frac{R_2 - R_1}{R_1} \times 100 = \text{Relative Improvement in Reliability}$$

repairable MTTF. It is to be noted from Table 11 that the relative improvement in reliability increases rapidly as the failure rate of the system element increases. For a lower failure rate of the system element, the system reliability approaches the upper limit (unity) by making the redundant system elements repairable.

To illustrate the advantages of incorporating the repair policy in system design, the following example is considered. Suppose that the system is required to have a reliability of, say, 0.90 or greater over the desired operating time of 100 hours. The failure rate of on-line element is 0.006, and that of off-line element is 0.003. Then, the reliability of a single redundant (two system elements) system without repair is estimated from Equation (5-11) to be 0.8332, whereas a doubly redundant (three system elements) system is 0.8902, which is still below the required level of 0.9. In this situation, the system designer may have the option of adding a third redundant unit or using the repair policy in order to meet the specified requirement. However, in this case, by making the single redundant system (two elements) repairable with E factor of 0.6, this requirement is satisfied with the reliability of 0.9188. And this improvement is obtained without adding much weight as compared with a case of adding a third element to the system. It is at this point that the designer makes a careful comparison between two approaches. If the weight and costs are of prime importance in the system design, obviously the design approach incorporating the repair policy is much preferable than the approach involving the straight duplications of the system elements.

Table 12. Relative Improvement in a Two-Element Repairable System's MTTF to a Non-Repairable System's MTTF where $\lambda_2 = \lambda_1/2$

E	μ = 0.01, λ_1 0.001				μ = 0.01, λ_1 = 0.002				μ = 1.0, λ_1 = 0.002				μ = 1.0, λ_1 = 0.01			
	M_1	M_2	ΔM_1	ΔM_2	M_1	M_2	ΔM_1	ΔM_2	M_1	M_2	ΔM_1	ΔM_2	M_1	M_2	ΔM_1	ΔM_2
0.1	1709	1733	1.03	1.04	853	864	1.02	1.04	857	870	1.03	1.04	171	174	1.03	1.04
0.2	1759	1814	1.06	1.09	875	900	1.05	1.08	884	917	1.06	1.10	177	183	1.06	1.10
0.4	1880	2048	1.13	1.23	929	1000	1.11	1.20	954	1055	1.15	1.27	190	210	1.14	1.26
0.6	2048	2467	1.23	1.48	1000	1167	1.20	1.40	1055	1331	1.27	1.60	210	264	1.26	1.59
0.8	2294	3444	1.38	2.07	1100	1500	1.30	1.80	1213	2154	1.46	2.58	241	420	1.45	2.52
1.0	2692	8333	1.62	5.00	1250	2500	1.50	3.00	1496	167500	1.80	2.01	296	6833	1.78	41
M_0	1667				833				833				167			

NOTE: M_0 = MTTF of Two-Element Non-Repairable System

M_1 = MTTF of Two-Element System where only a Primary Element is Repairable

M_2 = MTTF of Two-Element System where both Elements are Repairable

$\Delta M_1 = M_1/M_0$, $\Delta M_2 = M_2/M_0$

Concluding Remarks

When reliability allocations are made at various subsystem levels through the functional allocation method and the feasibility study discloses that the usual approaches of increasing reliability could not meet the allocated requirements, the system designer must then consider an approach in which the weak system elements are made repairable while in operation by an attending operator so that the resulting subsystem reliabilities should meet the allocated requirements.

In this chapter, a design approach is developed that is oriented toward the system designers who are faced with the problem of choosing on the proper combination between redundancy and the repair policy as a means of improving the desired reliability. This design approach is also helpful in that it allows the system designer to impose such logistic requirements as the level of spare parts and the availability of repair tools as well as the capability of the operator to perform the required repair work in order to achieve the allocated reliability goal of his system.

CHAPTER VI

CONCLUSIONS AND RECOMMENDATIONS

Conclusions

The system designer responsible for designing a reliable system for operational use must translate the overall system reliability requirement into numerical requirements at various subsystem levels. These allocations are performed by the use of an appropriate reliability allocation method. Because of the complexity in relating many interacting factors in the allocation process, the system reliability allocations are carried out in this research through two distinct methods--the functional allocation method, and the detailed allocation method. With the application of the functional allocation method, the system designer can assign a first set of the requirements to the functional equipment groups which comprise a given system. In developing this allocation method, considerations are given to such factors as the state-of-the-art, complexity, duty cycle as well as the applicational stresses, and use environment factors which influence unit reliability.

The functional allocation method developed here seems to have general applications due to the facts that (1) the required input data to the model are based only on the available information during the early design stages of a development program; (2) the reasoning employed in the allocation process is rather general so that the method can be applicable to various kinds of systems; and (3) the method tends to provide more realistic results as usage experiences in the field accumulate for a given system.

After the system reliability requirement has been allocated to the functional equipment groups within a system, the system designer must review and modify, if necessary, the allocated requirements as soon as the detailed feasibility study discloses the discrepancies between the allocated requirements and improvement feasibility. The most interesting case arises when there are several design approaches available for achieving the requirement including the conventional approach of employing redundancy for the weak units. The detailed allocation method developed for accomplishing the supplementary allocations takes explicitly account of this existence of several design approaches at each functional equipment level.

The allocation problem formulated in Equations (3-1) and (3-2), then, is to select the design alternative and the level of redundancy at each subsystem in such a way as to maximize the overall system reliability, when (a) there are several design approaches for achieving the required reliability, and (b) there also exist the constraints on system costs and weight for a given system.

An application of dynamic programming techniques to this allocation problem has resulted in a different formulation of the problem in Equations (3-12) and (3-13) in which the weight constraint is eliminated through the use of the Lagrange multiplier method. Whenever the reformulated allocation problem is solved for a given set of the Lagrange multiplier, a by-product of the computation of the optimal policy is a set of state values (the values of an eliminated constraint function in the optimal policy) that permit the sensitivity analysis in the given circumstances. In most cases of system reliability allocations, this sensitivity analysis is more interesting and useful than solving the problem with a particular value of the constraint. This kind of sensitivity analysis is practically impossible to obtain through the conventional approaches unless one has a fairly simple and explicit analytical solution to the problem.

The dynamic programming approach used to solve the problem is believed to be an efficient one since the computational scheme involved is simple, practical, easily implemented, and yet yields the exact solution to the problem. It is also believed that the computational scheme developed is applicable to a class of sequential decision problems involving two discrete decision variables at each stage with a set of constraints.

One important factor which is not incorporated in the development of the functional allocation method is the system cost which has a direct bearing in any development program. In Chapter IV, a reliability allocation method based on the concept of the system cost is presented in order to supplement the functional allocation method. This allocation method has several characteristics. It takes into account the prime cost elements that are relevant in the reliability allocation; and it makes clear distinction between the one-time investments and the recurring operating expenses which will not only permit an easy measurement of relevant cost elements but also actually minimize the possible errors in the cost estimations.

The final part of this study is concerned with the derivation of the reliability functions and the mean-time-to-failures for both the non-repairable redundant system and the repairable system. In certain situations, by making the redundant system repairable by an attending operator, the system designer can increase reliability of his system without adding much system weight as compared with a case of the straight duplications of critical system elements. In general, the system designer desires to find the relative improvement in system reliability or mean-time-to-failure of a repairable system over its non-repairable counterpart.

The reliability function and the equation for mean-time-to-failure developed make it possible to estimate these relative improvements. The choice between two design approaches--one incorporating the repair policy which allows instant repairs while in operation, and that of adding one or more redundant units to the weak elements can then be made in certain design situations with due considerations given to such factors as increased system weights or system costs.

Recommendations

In the course of carrying out this research, it has found some areas of the future research that would be fruitful in further refinement of the allocation method. There are two areas for which additional research may follow:

1. Testing of basic assumptions made. In developing the functional allocation method, the underlying failure distribution is assumed to be the exponential form, and that the failure probabilities

of the allocation units are independent since the accurate information on the failure process of a newly developed system is not available at the early stages of a development program. Therefore, it was assumed that the optimal allocation derived under these assumptions is minimally sensitive. Admittedly, not all system reliability allocation problems can be handled in this way. Hence, one extension of this research involves formulating new assumptions about the failure process and making necessary comparisons between the two methods.

2. Testing of the validity of assumed cost functions. One of the crucial tasks in applying the allocation method based on the concept of the system costs is the determination from historical data of useful relationships between the relevant system costs and the allocated failure rate. Because there are insufficient data for estimating this cost function in practice, the form of the cost function is largely judged by the engineering experiences. Therefore, the propositions made in the derivation of this cost function may be further verified by the actual observations of real world cases.

APPENDIX

COMPUTER PROGRAMMING PROCEDURE

IN ALGOL

```
                                                                                    0000
                                                          START OF SEGMENT ********** 0002
INTEGER        N, CMAX, B, I, J, E, G, BETA, Z, ZHAT, MBAR, YY, K, T ;              0000
INTEGER        SUM, H, TOTALCOST, TOTALWEIGHT, Q ,R        ;                        0000
REAL           V, D, LARGEST, GAIN, FIRST,RELIABILITY,TT   ;                        0000
INTEGER ARRAY AT,MHAT[0:22],C,W[0:22,0:4 ], M,Y[0:22,0:150]    ,                    0000
               MSTAR,  YSTAR[0:22] , WC, SS[0:150]  ;                               0006
REAL ARRAY     MAX[0:15], P[0:22,0:4 ], F,S[0:150]  ;                               0010
LABEL          L2    ;                                                              0016
FILE IN        CARD(2,10) ;  FILE OUT  LINE 4(2,15) ;                               0016
LIST           LT1(N, CMAX, FOR I ← 1 STEP 1 UNTIL N  DO AT[I]) ;                   0023
LIST           LT2(FOR I← 1 STEP 1 UNTIL N DO FOR J← 1 STEP 1 UNTIL                 0034
                                     AT[I]  DO P[I,J]) ;                            0037
LIST           LT3(FOR I← 1 STEP 1 UNTIL N DO FOR J← 1 STEP 1 UNTIL                 0049
                                     AT[I]  DO C[I,J]) ;                            0052
LIST           LT4(FOR I← 1 STEP 1 UNTIL N DO FOR J← 1 STEP 1 UNTIL                 0064
                                     AT[I]  DO W[I,J]) ;                            0067
FORMAT OUT     FOM1("LAMBDA  =",X2,F7.4,X5,"I =",X2,I4),                            0079
                                                          START OF SEGMENT ********** 0003
               FOM2(X11,"COST",X4,"F(I,B)",X6,"M(I,B)",X4,"Y(I,B)",                 0079
               X4,"WC(I,B)"),                                                       0079
               FOM3(I13,X5,F8.4,X5,I5,X5,I5,X5,I5 ),                                0079
               FOM4(X11,"SUBSYSTEM NO",X6,"REDUNDANCY",X8,"DESIGN NO"),             0079
               FOM5(3(I17))   ,                                                     0079
               FOM6(X11,"RELIABILITY   =",F10.4//X7,"TOTAL COST   =",I5,            0079
               X5,"TOTAL WEIGHT    =",I5) ,                                         0079
               FOM7(X11,"TIME  =",F5.1)    ;                                        0079
                                                          0003 IS 0082 LONG, NEXT SEG 0002
               READ(CARD,/ ,LT1)     ;   READ(CARD,/ ,LT2) ;                        0079
               READ(CARD,/ ,LT3);   READ(CARD,/ ,LT4);                              0086
               CLOSE(CARD, RELEASE)    ;                                            0093
               FOR  V  ← 0.0008  DO                                                 0095
```

```
        FOR  V  ←  0.0009  DO
    BEGIN
        FOR B ←0 STEP 1 UNTIL CMAX DO BEGIN F[B] ← 1    ;
        WC[B]  ← 0   END   ;
        TT  ←  TIME(1)   ;
        FOR I ← 1 STEP 1 UNTIL N DO
    BEGIN
        FOR B ← 0 STEP 1 UNTIL CMAX DO
    BEGIN
        FOR J ← 1 STEP 1 UNTIL AT[I] DO
    BEGIN
        D ← P[I,J]  ;  F ← C[I,J]  ;  G ← W[I,J]  ;
        LARGEST ← -2.0 ;
        BETA ← ENTIER(B/E)  ;
        IF BETA > 0 THEN Z ← 1 ELSE Z ← 0  ;
L2:     GAIN ← (1-(1-D)↑Z) × EXP(-V×Z×G) ×F[B-Z×E]    ;
        IF GAIN > LARGEST THEN BEGIN LARGEST ← GAIN ;
        ZHAT ← Z  END ;
        Z ← Z + 1  ;
        IF Z > BETA THEN BEGIN MAX[J] ← LARGEST  ;
        MHAT[J] ← ZHAT END ELSE GO TO L2
    END;
        FIRST ← MAX[1]  ;
        MBAR ←MHAT[1]  ;
        YY ← 1  ;
        FOR K ← 2 STEP 1 UNTIL AT[I] DO
        IF MAX[K] > FIRST THEN BEGIN FIRST ← MAX[K]   ;
        MBAR ← MHAT[K] ;  YY ← K  END  ;
        S[B]  ← FIRST   ;
        M[I,B] ← MBAR  ;
        Y[I,B] ← YY        ;
        SS[B] ← MBAR × W[I,YY]  + WC[B-MBAR × C[I,YY]]
```

```
        END;                                                              016
            FOR R ← 0 STEP 1 UNTIL CMAX DO BEGIN F[R] ← S[R] ;            017
            WC[R] ← SS[R]   END  ;                                        017
        END;                                                              017
            WRITE(LINE,FOM1,V,N)    ;                                     01
            WRITE(LINE,FOM2)  ;                                           01
            FOR T ← 0 STEP 1 UNTIL CMAX DO                                0193
            WRITE(LINE,FOM3,T,F[T],M[N,T],Y[N,T],WC[T])      ;            0194
            SUM  ← CMAX   ;                                               0211
            FOR H ← N  STEP -1  UNTIL 1  DO BEGIN                         0212
            MSTAR[H] ← M[H,SUM]   ;                                       0213
            YSTAR[H] ← Y[H,SUM]   ;                                       0215
            SUM ← SUM-MSTAR[H] × C[H,YSTAR[H]]   END   ;                  0217
            TOTALCOST  ← CMAX - SUM    ;                                  0223
            TOTALWEIGHT   ← WC[CMAX]   ;                                  0224
            RELIABILITY  ← F[CMAX] × EXP(V × TOTALWEIGHT  )  ;            0225
            WRITE(LINE[PAGE])  ;                                          0227
            WRITE(LINE, FOM4)  ;                                          0230
            FOR  Q ← 1 STEP 1 UNTIL N  DO                                 0233
            WRITE(LINE,FOM5,Q,MSTAR[Q], YSTAR[Q])   ;                     0235
            WRITE(LINE,FOM6,RELIABILITY, TOTALCOST, TOTALWEIGHT)  ;       0247
            TT  ←(TIME(1) - TT) / 60  ;                                   0258
            WRITE(LINE,FOM7,TT)                                           0260
        END;                                                              0265
            END.                                                          0269
                                                0002 IS 0272 LONG, NEXT SEG 0001
EXP IS SEGMENT NUMBER 0004,PRT ADDRESS IS 0110
LN IS SEGMENT NUMBER 0005,PRT ADDRESS IS 0107
OUTPUT(W) IS SEGMENT NUMBER 0006,PRT ADDRESS IS 0113
BLOCK CONTROL IS SEGMENT NUMBER 0007,PRT ADDRESS IS 0005
INPUT(W) IS SEGMENT NUMBER 0008,PRT ADDRESS IS 0106
X TO THE I IS SEGMENT NUMBER 0009,PRT ADDRESS IS 0111
```

BIBLIOGRAPHY

BIBLIOGRAPHY

1. Advisory Group on Reliability of Electronic Equipment, *Reliability of Military Electronic Equipment.* Washington, D. C.: U. S. Government Printing Office (June, 1957).

2. Albert, A., *A Measure of the Effort Required to Increase Reliability,* Technical Report No. 43, Applied Mathematics and Statistics Laboratory, Stanford University (November, 1958).

3. Albert, A., and F. Proschan, *Increased Reliability with Minimum Effort,* Technical Report No. 50, Applied Mathematics and Statistics Laboratory, Stanford University (October, 1959).

4. ARINC Research Corporation, *Improved Techniques for Design State Prediction,* Air Force Reliability Assurance Program Progress Report, Vol. I, No. 2, Contract AF 33(600) - 38438, Washington, D. C.: ARINC Research Corporation.

5. ARINC Research Corporation, *The Allocation of System Reliability,* Technical Documentary Report, No. ASD-TDR-62-20, Vol. I, Washington, D. C.: ARINC Research Corporation.

6. Barbara, Fox, "Total Annual Cost, A Reliability Criterion," *Proceedings of the Tenth National Symposium on Reliability and Quality Control* (January, 1964), pp. 266-273.

7. Barlow, R. E., and F. Proschan, *Mathematical Theory of Reliability.* New York: John Wiley & Sons, Inc., 1965.

8. Barlow, R. E., and L. C. Hunter, "Criteria for Determining Optimum Redundancy," *IRE Transactions on Reliability and Quality Control* (April, 1960), pp. 73-77.

9. Barlow, R. E., and L. C. Hunter, "System Efficiency and Reliability," *Technometrics,* Vol. 2 (1960), pp. 43-53.

10. Barlow, R. E., L. C. Hunter, and F. Proschan, "Optimum Checking Procedures," *Proceedings of the Seventh National Symposium on Reliability and Quality Control* (January, 1961), pp. 485-495.

11. Barlow, R. E., and L. C. Hunter, *Mathematical Models for System Reliability,* Sylvania Electric Defense Laboratory Report, Prepared for U. S. Army Signal Engineering Laboratories, Contract DA-36-039, SC 78281, AD 228131 (August, 1959).

12. Barlow, R. E., L. C. Hunter, and F. Proschan, "Optimum Redundancy when Components are Subject to Two Kinds of Failures," *Journal of the Society for Industrial and Applied Mathematics*, Vol. 11, No. 1 (March, 1963), pp. 64-73.

13. Bazovsky, Igor, *Reliability, Theory and Practice*. Englewood Cliffs, N. J.: Prentice-Hall Inc., 1961.

14. Bellman, Richard E., "Dynamic Programming and Lagrange Multipliers," *Proceedings of National Academy of Science, U.S.A.*, Vol. 42 (1956), pp. 767-769.

15. Bellman, Richard E., *Dynamic Programming*. Princeton, N. J.: Princeton University Press, 1957.

16. Bellman, Richard E., and Stuart E. Dreyfus, *Applied Dynamic Programming*. Princeton, N. J.: Princeton University Press, 1962.

17. Bellman, Richard E., and Stuart E. Dreyfus, "Dynamic Programming and the Reliability of Multicomponent Devices," *Journal of the Operations Research Society of America*, Vol. 6 (1958), pp. 200-206.

18. Bellman, Richard E., "Some Applications of the Theory of Dynamic Programming," *Journal of the Operations Research Society of America*, Vol. 2 (1954), pp. 275-288.

19. Birnfaum, Z. W., and S. C. Saunders, "A Statistical Model for Life Length of Materials," *Journal of the American Statistical Association*, Vol. 53, No. 281 (1959), pp. 151-160.

20. Black G., and F. Proschan, "On Optimal Redundancy," *Journal of the Operations Research Society of America*, Vol. 7 (1957), pp. 581-588.

21. Black, G., and F. Proschan, "Spare Part Kits at Minimum Cost," *Proceedings of the Fifth National Symposium on Reliability and Quality Control* (January, 1959), pp. 281-289.

22. Bollman, T. H., "Instructions and Data for Failure and Prediction," *Bell Telephone Laboratories Report* (October, 1957).

23. Breipohl, A. M., "A Unique Allocation of Required Component Reliability," *Proceedings of the Seventh National Symposium on Reliability and Quality Control* (January, 1961), pp. 189-202.

24. Bureau of Naval Weapons, *Reliability Engineering Handbook*, NAVWEPS 00-65-502. Washington, D. C.: U. S. Government Printing Office (June, 1964).

25. Cafars, J. A., and H. D. Voegtlen, "The Measurement and Specification of Product Abilities," *Industrial Quality Control*, Vol. XVIII, No. 9 (1962), pp. 20-26.

26. Carhart, R. R., and G. R. Herd, "A Simple Cost Model for Optimizing Reliability," *Reliability of Military Electronic Equipment*, Appendix F, AGREE Report (June, 1957).

27. Carhart, R. R., *A Survey of the Current Status of the Electronic Reliability Problems*, The RAND Research Memorandum RM-1131, the RAND Corporation (August, 1953).

28. Churchman, C. W., R. L. Ackoff, and E. L. Arnoff, *Introduction to Operations Research.* New York: John Wiley & Sons, Inc., 1957.

29. Col. V. J. Bracha, *Analysis of Reliability Management in Defense Industries*, U.S.A.F. Ballestic Systems Division (BSD-62-48).

30. Coleman, J. J., and I. J. Abrams, "Mathematical Model for Operational Readiness," *Journal of the Operations Research Society of America*, Vol. 10 (1962), pp. 126-138.

31. Cox, D. R., and W. L. Smith, "On the Superposition of Renewal Processes," *Biometrica*, Vol. 41 (1954), pp. 91-99.

32. Cox, D. R., "The Analysis of Experimentally Distributed Life Times with Two Types of Failures," *Journal of the Royal Statistical Society*, Series B, No. 21 (1959), pp. 411-421.

33. Davis, D. J., "Analysis of Some Failure Data," *Journal of the American Statistical Association*, Vol. 47, No. 258 (1952), pp. 113-150.

34. Dorn, W. S., "Non-Linear Programming," *Management Science*, Vol. 9, pp. 171-208.

35. Dorn, W. S., "Lagrange Multipliers and Inequalities," *Journal of the Operations Research Society of America*, Vol. 9, No. 1 (1961), pp. 95-104.

36. Drenick, R. F., "The Failure Law of Complex Equipment," *Journal of the Society for Industrial and Applied Mathematics*, Vol. 8, No. 4 (1960), pp. 680-690.

37. Drenick, R. F., "Mathematical Aspects of the Reliability Problems," *Journal of the Society for Industrial and Applied Mathematics*, Vol. 8 (1960), pp. 125-149.

38. Electronic Industries Association, "Determination of Permissible Component Part Failure Rates," *Electronic Application Review*, Vol. 4, No. 1 (September, 1956), pp. 10-12.

39. Epstein, B., and M. Sobel, "Some Theorem Relevant to the Life Testing from an Exponential Distribution," *The Annals of Mathematical Statistics*, Vol. 25 (1954), pp. 373-381.

40. Epstein, B., "The Exponential Distribution and its Role in Life Testing," *Industrial Quality Control*, Vol. 15, No. 6 (1958), pp. 4-9.

41. Epstein, B., and M. Sobel, "Life Testing," *Journal of the American Statistical Association*, Vol. 48, No. 263 (1953), pp. 486-502.

42. Everett, H., "Generalized Lagrange Multiplier for Solving Problems of Optimum Allocation of Resources," *Journal of the Operations Research Society of America*, Vol. 2 (1963), pp. 399-415.

43. Feller, William, *An Introduction to Probability Theory and its Applications*, Vol. I; New York: John Wiley & Sons Inc., 1953.

44. Flehinger, B. J., "A General Model for the Reliability Analysis of System under Various Preventive Maintenance Policies," *Annals of Mathematical Statistics*, Vol. 33 (1962), p. 137-156.

45. Flehinger, B. J., "Reliability Improvement Through Redundancy at Various System Levels," *IBM Journal of Research and Development*, Vol. 2, No. 2 (April, 1958), pp. 148-158.

46. Frederic, H. E., "A Reliability Allocation Technique," *Proceedings of the Fourth National Symposium on Reliability and Quality Control* (January, 1958), pp. 314-317.

47. Garbarino, H. L., "Selection of Reliability Levels in Equipment Design," *IRE Transactions on Industrial Electronics*, IE-5 (April, 1958), pp. 76-81.

48. Geisler, M. A., and H. W. Karr, "The Design of Military Supply Tables for Spare Parts," *Journal of the Operations Research Society of America*, Vol. 4, No. 4 (1956), pp. 431-442.

49. Gordon, R., "Optimum Component Redundancy for Maximum System Reliability," *Journal of the Operations Research Society of America*, Vol. 5, No. 4 (1957), pp. 229-243.

50. Gourary, M. H., "A Simple Rule for the Consolidation of Allowable List," *Naval Research Logistics Quarterly*, Vol. 5, No. 5 (1958), pp. 1-15.

51. Green, A. W., and R. C. Horne, *Maintainability of Shipboard Electronic Systems*, ARINC Research Corporation Publication No. 118-4-228, Washington, D. C.: ARINC Research Corporation, 1960.

52. Hadley, G., *Non-Linear and Dynamic Programming.* Reading, Mass.: Addison-Wesley Publishing Company, Inc., 1964.

53. Hees, R. N. V., and H. W. Meerendonk, "Optimal Reliability of Parallel Multicomponent Systems," *Operational Research Quarterly,* Vol. 12, No. 1 (1961), pp. 19-26.

54. Hershey, J. H., "Reliability and Maintainability of Military Electronic Equipment," *Proceedings of the Third Annual Signal Maintenance Symposium,* Fort Monmouth, New Jersey (April, 1959).

55. Hitch, C. J., and R. N. McKean, *The Economics of Defense in the Nuclear Age.* Cambridge, Mass.: Harvard University Press, 1960.

56. Johnston, D. E., and D. T. McRuer, *A Summary of Component Failure Rate and Weighting Function Data and their Use in System's Reliability Design,* Wright Air Development Center WADC Tr.57-668 (December, 1957).

57. Johnston, D. E., K. A. Ferrick, and E. L. Alexander, *A Compilation of Component Field Reliability Data Useful in Preliminary Design,* WADD Technical Report 60-330 (September, 1961).

58. Kahn, H., and I. Mann, *Techniques of System Analysis,* The RAND Corporation Research Memorandum RM-1829-1 (June, 1957).

59. Kao, J. H. K., "A Summary of Some New Techniques on Failure Analysis," *Proceeding of the Sixth National Symposium on Reliability and Quality Control* (January, 1960).

60. Kemeny, J. G., and J. L. Snell, *Finite Markov Chain.* Princeton, N. J.: D. Van Nostrand Company, Inc., 1960.

61. Kettelle, J. D., "Least Cost Allocation of Reliability Investment," *Journal of the Operations Research Society of America,* Vol. 10 (1962), pp. 249-265.

62. Khintchine, A. Ya., "Mathematisches Uber die Erwartung von einen Offentlicher Schalter," *Matem. Sbornik* (1932).

63. Kuhn, H. W., and A. W. Tucker, "Non-Linear Programming," *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability* (December, 1954), pp. 115-128.

64. Liittschwagner, J. M., "Dynamic Programming in the Solution of a Multistage Reliability Problem," *The Journal of Industrial Engineering* (July-August, 1964), pp. 168-175.

65. Loyd, D. K., and M. Lipow, *Reliability: Management, Methods, and Mathematics.* Englewood Cliffs, N. J.: Prentice-Hall Inc., 1962.

66. Madison, R. L., *Effects of Maintenance on System Reliability*, ARINC Research Corporation Publication 101-16-144, Washington, D. C.: ARINC Research Corporation, 1959.

67. Mayne, A. J., "Some Reliability Models of Production Lines with Special Reference to Computer Operation and Scheduling," *Operational Research Quarterly*, Vol. 11 (1960), pp. 16-30.

68. Mine, H., "Reliability of Physical System," *IRE Transactions PGIT*, IT-5, Special Supplement (1959).

69. Mitten, L. G., and G. L. Nemhauser, "Multistage Optimization," *Chemical Engineering Progress*, Vol. 59, No. 1, pp. 52-60.

70. Morrison, D. F., and H. A. David, "The Life Distribution and Reliability of a System with Spare Components," *Annals of Mathematical Statistics*, Vol. 31, No. 4 (1960), pp. 1084-1094.

71. Moskowitz, F., and J. B. McLean, "Some Reliability Aspects of System Design," *IRE Transactions on Reliability and Quality Control*, PGRQC-8 (1956), pp. 50-59.

72. Moskowitz, F., "The Statistical Analysis of Redundant Systems," *IRE International Convention Record*, Part 6 (March, 1960), pp. 78-89.

73. Mayne, A. J., K. L. Wong, and H. M. Gordy, *Reliability Engineering for Electronic Systems*. New York: John Wiley & Sons, Inc., 1964.

74. Noah, J. W., *Identifying and Estimating R & D Costs*, The RAND Corporation Research Memorandum RM-3067-PR (May, 1962).

75. Novick, D., *Weapon System Cost Methodology*, The RAND Corporation Report R-287 (February, 1956).

76. Novick, D., *System and Total Force Cost Analysis*, The RAND Corporation Research Memorandum RM-2695 (April, 1961).

77. Novick, D., "Weapon System Cost Analysis," *IRE Transactions on Engineering Management*, EM-3, Vol. 1, No. 3 (1956), pp. 67-73.

78. Palm, Docent C., "Arbetskraftens Fordelning vid betjaning av automatskiner," *Industritidningen Norden*, Vol. 75 (1947), pp. 75-80, 90-94.

79. Proschan, F., "Optimal System Supply," *Naval Research Logistics Quarterly*, Vol. 7, No. 4 (1961), pp. 609-646.

80. Proschan, F., *Minimum Cost for Spares when Component Failure is not Necessary Exponential*, Report EDL-M185, Electronic Defense Laboratories, Mountain View, California (1959).

81. Proschan, F., and T. A. Bray, *Optimum Redundancy under Multiple Constraints*, Boeing Scientific Research Laboratory Report DI-82-0253 (1963).

82. Radio Corporation of America, *Reliability Stress Analysis for Electronic Equipment*, RCA Technical Report 59-46-1 (January, 1959).

83. RCA Service Company, *The Prediction and Measurement of Air Force Ground Electronic Equipment Reliability*, Final Engineering Report, Contract No. AF 30(602) - 1623 RADCTN 58-307 (August, 1958).

84. Roberts, N. H., *Mathematical Methods in Reliability Engineering*. New York: McGraw-Hill Book Company, 1964.

85. Rosenblatt, J. R., "On Prediction of System Performance from Information on Component Performance," *Proceedings of Western Joint Computer Conference* (February, 1957), pp. 85-94.

86. Ryerson, C. M., "Reliability Testing Theory Based on the Poisson Distribution," *Proceedings of the Fourth National Symposium on Reliability and Quality Control* (January, 1958), pp. 3-18.

87. Sandler, G. H., *System Reliability Engineering*. Englewood Cliffs, N. J.: Prentice-Hall Inc., 1963.

88. Sasaki, M., "A Simplified Method of Obtaining Highest System Reliability," *Proceedings of the Eighth National Symposium on Reliability and Quality Control* (January, 1962) pp. 489-502.

89. Simon, H. A., "Dynamic Programming under Uncertainty with a Quadratic Function," *Econometrica*, Vol. 24 (1956), pp. 74-81.

90. Stoller, D. S., "A Failure Model for Equipment Undergoing Complex Operations," *Journal of the Operations Research Society of America*, Vol. 6, No. 5 (1958), pp. 723-728.

91. Truelove, A. J., "Mathematical Models for Optimizing Strategic Reliability and for Minimizing Costs," *Proceedings of the Sixth Joint Military Industry Guided Missile Reliability Symposium*, Vol. 2 (February, 1960), pp. 87-108.

92. Truelove, A. J., "Strategic Reliability and Preventive Maintenance," *Journal of the Operations Research Society of America*, Vol. 9 (1961), pp. 22-29.

93. United States Department of Defense, *Sampling Procedures and Tables for Life and Reliability Testing* (Department of Defense Handbook H 108). Washington, D. C.: U. S. Government Printing Office, 1960.

94. United States Department of Defense, *Reliability Stress and Failure Rate Data for Electronic Equipment* (MIL-HDBK-217). Washington, D. C.: U. S. Government Printing Office, 1962.

95. United States Department of Defense, *Climatic Extremes for Military Equipment* (MIL-STD-210A) Washington, D. C.: U. S. Government Printing Office.

96. United States Department of Defense, *Reliability of Military Electronic Equipment* (MIL-STD-441). Washington, D. C.: U. S. Government Printing Office, 1958.

97. United States Department of Defense, *Reliability Prediction* (MIL-STD-756A). Washington, D. C.: U. S. Government Printing Office, 1961.

98. United States Department of Defense, *Reliability of Production Electronic Equipment* (MIL-R-19610). Washington, D. C.: U. S. Government Printing Office, 1956.

99. United States Department of Defense, *Reliability Requirements for Design of Electronic Equipment or System* (MIL-R-22256). Washington, D. C.: U. S. Government Printing Office, 1959.

100. United States Department of Defense, *Reliability Requirements for Shipboard and Ground Electronic Equipment* (MIL-R-22732). Washington, D. C.: U. S. Government Printing Office, 1961.

101. United States Department of Defense, *Reliability Index Determination for Avionic Equipment* (MIL-R-22973). Washington, D. C.: U. S. Government Printing Office.

102. United States Department of Defense, *Reliability Assurance for Production Acceptance of Avionic Equipment* (MIL-R-23094). Washington, D. C.: U. S. Government Printing Office.

103. United States Department of Defense, *Reliability Assurance Program for Electronic Equipment* (MIL-25717A). Washington, D. C.: U. S. Government Printing Office, 1959.

104. United States Department of Defense, *Reliability Requirements for Production Ground Electronic Equipment* (MIL-R-26474). Washington, D. C.: U. S. Government Printing Office, 1959.

105. United States Department of Defense, *Reliability Requirements for Development of Electronic Subsystems or Equipment* (MIL-R-26484). Washington, D. C.: U. S. Government Printing Office, 1958.

106. United States Department of Defense, *Reliability and Longevity Requirement, Electronic Equipment, General Specifications for* (MIL-R-26667A). Washington, D. C.: U. S. Government Printing Office, 1959.

107. United States Department of Defense, *Reliability Requirements for Weapon System, General Specification* (MIL-R-26674). Washington, D. C.: U. S. Government Printing Office, 1959.

108. United States Department of Defense, *Reliability Requirements for Development of Ground Electronic Equipment* (MIL-R-27070). Washington, D. C.: U. S. Government Printing Office, 1960.

109. United States Department of Defense, *Reliability Program Requirements for Aerospace System, Subsystems, and Equipment* (MIL-R-27542). Washington, D. C.: U. S. Government Printing Office, 1961.

110. United States National Aeronautics and Space Administration, *Reliability Program Provisions for Space System Contractors* (NPC 250-1). Washington, D. C.: U. S. Government Printing Office, 1963.

111. Von Alven, W. H., *Reliability Engineering,* Englewood Cliffs, N. J.: Prentice-Hall Inc., 1964.

112. Wagner, D. H., "Numerical Reliability Requirements," *Proceedings of the Third National Symposium on Reliability and Quality Control* (January, 1957), pp. 154-158.

113. Weiss, G. H., "The Reliability of Components Exhibiting Cumulative Damage Effects," *Technometrics,* Vol. 3 (1961), pp. 413-422.

114. Wilde, D. J., "Differential Calculus in Non-Linear Programming," *Journal of the Operations Research Society of America,* Vol. 10 (1962), pp. 764-773.

115. Zelen, M., and M. C. Dannemiller, "The Robustness of Life Testing Procedures Derived from the Exponential Distribution," *Technometrics,* Vol. 3, No. 1, pp. 29-49.

116. Zelen, M., *Statistical Theory of Reliability.* Madison, Wisconsin: The University of Wisconsin Press, 1963.

VITA

Nam Kee Lee was born in Seoul, Korea, on August 13, 1930, the son of Dr. and Mrs. Byung Wee Lee. He attended public school in Seoul, Korea, where he was graduated from high school in May, 1949, with the highest honor.

In September of 1949, the author entered Seoul National University where he undertook a program of study in mechanical engineering. Upon the outbreak of the Korean War in 1950, he joined the Korean Army as a first lieutenant and served for the next four years in various combat duties. He was awarded the Bronze Star with Oak-leaf Cluster for combat leadership and was discharged from active duty in November, 1954, with the rank of major.

While serving in the Army, the author took the extension course of Seoul National University and fulfilled the requirements for the undergraduate degree. He was graduated in March, 1954, with the degree of Bachelor of Science in Mechanical Engineering. Beginning September 1954, the author enrolled at Seoul National University to pursue graduate study and receive the degree Master of Science in Mechanical Engineering in March, 1956.

During May of 1956, the author was employed by the Korea Fertilizer Corporation as a project engineer and supervised the entire construction of its Urea fertilizer plant in Chungju, Korea. In September, 1957, he was assigned to Switzerland to coordinate industrial projects for the Korean Government.

Returning to Korea in December, 1958, the author resumed his duty with the Korea Fertilizer Corporation as Chief Engineer at the Chungju Plant. In June of 1961, he enrolled at the National Institute of Public Administration in Seoul, Korea, to carry a program of study in public administration. He completed the requirements of the Institute in May, 1962 and was awarded the highest honor from the President of the Republic of Korea.

To pursue further graduate study, the author enrolled at the University of Houston in September, 1962, and received the degree Master of Science in Industrial Engineering in June, 1964. In September of 1964, he enrolled in the Ph.D. program of the School of Industrial Engineering at the Georgia Institute of Technology. He is a member of Alpha Pi Mu honorary society. The author is married and has two sons.